

# data analitik buku

*by* Rossi Passarella

---

**Submission date:** 15-Nov-2021 12:36PM (UTC+0700)

**Submission ID:** 1703081890

**File name:** Buku\_-\_Daffa\_final.pdf (3.67M)

**Word count:** 20802

**Character count:** 133756

Book Series : Aircraft Accidents - Incidents

Rossi Passarella  
M. Daffa Badran Thoriq  
Siti Nurmaini

# Data Analysis: The Effect of Changes that Occurred After ICAO Standardization on World Civil Aviation Accidents

BAAA Database From 1918 - 2020



13  
UNSRI  
PRESS

**BOOK SERIES: AIRCRAFT ACCIDENTS-INCIDENTS**

**DATA ANALYSIS:**

**THE EFFECT OF CHANGES THAT OCCURRED  
AFTER ICAO STANDARDIZATION ON  
WORLD CIVIL AVIATIONS ACCIDENTS**

BAAA DATABASE FROM 1998-2020

ROSSI PASSARELLA  
M.DAFFA BADRAN THORIQ  
SITI NURMAINI



*This Page Intentionally Left Blank*



**Sanksi pelanggaran Pasal 72  
Undang-undang Nomor 19 Tahun 2002  
Tentang Perubahan atas Undang-undang Nomor 12 Tahun 1997  
Pasal 44 Tentang Hak Cipta**

1. Barang siapa dengan sengaja dan tanpa hak melakukan perbuatan sebagaimana dimaksud dalam Pasal 2 ayat (1) atau pasal 49 ayat (1) dan ayat (2) dipidana penjara masing-masing paling singkat 1 (satu) bulan dan/atau denda paling sedikit Rp. 1.000.000,00 (satu juta rupiah), atau pidana penjara paling lama 7 (tujuh) tahun dan/atau denda paling banyak Rp. 5.000.000.000,00 (lima miliar rupiah)
2. Barang siapa dengan sengaja menyiarkan, memamerkan, mengedarkan, atau menjual kepada umum suatu ciptaan atau barang hasil pelanggaran hak cipta atau hak terkait, sebagaimana dimaksud ayat (1) dipidana dengan pidana penjara paling lama 5 (lima) tahun dan/atau denda paling banyak Rp. 500.000.000,00 (lima ratus juta rupiah)

*This Page Intentionally Left Blank*

## PREFACE

Bismillahirrahmanir Rahim,

Praise and gratitude the author prays to Allah SWT, who has given strength and perseverance so the book titled "The Effect of Changes that Occurred After ICAO Standardization on World Civil Aviation Accidents" can finally be completed.

This book is specially prepared for readers to understand the meaning of international commercial aviation, the role of ICAO as an international civil aviation safety organization for international aviation traffic, processes and techniques in data analytics, and how to analyze international commercial airplane crash data using descriptive statistics and exploratory data analysis (EDA) methods.

The author would like to thank various parties who have assisted in the preparation of this book. The author also feels that this book is far from perfect. Therefore, suggestions and constructive criticism are undoubtedly welcome. Hopefully, this book can be helpful for anyone who wants to learn the steps in analyzing other aircraft accident data using descriptive statistics and EDA methods.

Palembang, Agustus 2021

*This Page Intentionally Left Blank*

## TABLE OF CONTENTS

BOOK SERIES: AIRCRAFT ACCIDENTS-INCIDENTS.....	2
DATA ANALYSIS:.....	2
PREFACE .....	6
TABLE OF CONTENTS.....	8
LIST OF FIGURES.....	12
LIST OF TABLES.....	16
1 CHAPTER 1 INTRODUCTION.....	20
1.1 Background of Study .....	20
1.2 Problem formulation.....	23
1.3 Scope of the problems.....	24
1.4 The purposes of this study.....	24
1.5 The benefits of this study .....	25
2 CHAPTER 2 LITERATURE REVIEW.....	29
2.1 Related Studies .....	29
2.2 History of ICAO.....	37
2.3 Aviation Accident .....	39
2.4 Civil Aviation .....	41
2.4.1 Commercial Aviation.....	43
2.4.2 General Aviation.....	44
2.5 Data Analysis.....	45
2.5.1 Determine Analysis Objectives .....	46

	84		
2.5.2		Data Collection .....	47
2.5.3		Data Preprocessing .....	48
2.5.4		Descriptive Statistics.....	51
2.5.5		Exploratory Data Analysis (EDA).....	71
3		CHAPTER 3 RESEARCH METHODOLOGY.....	76
3.1		Determine Analysis Objectives and Create Hypothesis.	79
3.2		Data Collection.....	80
3.3		Data Preprocessing.....	81
3.3.1		Load Data .....	82
3.3.2		Renaming The Variable's Column Name .....	82
3.3.3		Rename The Variable's Name.....	83
3.3.4		Remove The Accidents Data of Non-Commercial Flight	84
3.3.5		Remove The Data which Contain Unknown Causes of Accidents.....	85
3.3.6		Variables Selection.....	86
3.4		Descriptive Statistics.....	86
3.4.1	83	Quantitative Data .....	87
3.4.2		Qualitative Data.....	93
3.5		Exploratory Data Analysis (EDA).....	94
4	92	CHAPTER 4 RESULTS AND DISCUSSIONS .....	99
4.1		Descriptive Statistics.....	99
4.1.1		The Variables with Quantitative Data Type .....	99
4.1.2		The Variables with Qualitative Data Type .....	121

4.2	Exploratory Data Analysis.....	129
6		
4.2.1	Number of Accidents and Fatalities per Year.....	130
6		
4.2.2	Number of Accidents and Fatalities per Year by Flight Types.....	135
6		
4.2.3	Number of Accidents and Fatalities per Year by Probable Causes.....	151
4.2.4	Summary Analysis.....	173
5	CHAPTER 5 CONCLUSIONS.....	182
6	CHAPTER 6 CLOSING STATEMENT.....	188
	GLOSARIUM.....	189
	INDEX.....	191
8	BIBLIOGRAPHY.....	195

*This Page Intentionally Left Blank*



## LIST OF FIGURES

<b>Gambar 2.1</b> Klasifikasi penerbangan sipil (civil aviation) berdasarkan ICAO .....	43
<b>Gambar 2.2</b> Jenis distribusi suatu data .....	56
<b>Gambar 2.3</b> Persebaran setiap nilai dari titik pusat distribusi data pada kondisi (a). Low dispersion; (b). High dispersion .....	62
<b>Gambar 2.4</b> Distribusi data berdasarkan nilai kurtosis .....	70
<b>Gambar 3.1</b> Tahap Penulisan .....	78
<b>Gambar 4.1</b> Visualisasi pengukuran frekuensi kemunculan pada variabel years .....	101
<b>Gambar 4.2</b> Visualisasi pengukuran frekuensi kemunculan pada variabel total fatalities .....	103
<b>Gambar 4.3</b> Visualisasi pemusatan data pada variabel years .....	106
<b>Gambar 4.4</b> Visualisasi pemusatan data pada variabel total fatalities .....	109
<b>Gambar 4.5</b> Visualisasi persebaran titik data dari titik pusat distribusi data pada variabel years .....	113
<b>Gambar 4.6</b> Visualisasi persebaran titik data dari titik pusat distribusi data pada variabel total fatalities.....	115
<b>Gambar 4.7</b> Visualisasi frekuensi kemunculan masing-masing aktivitas penerbangan komersial.....	123
<b>Gambar 4.8</b> Visualisasi persentase masing-masing aktivitas penerbangan komersial pada data.....	124
<b>Gambar 4.9</b> Frequency Distribution pada variabel Probable Causes .....	126
<b>Gambar 4.10</b> Persentase kategori penyebab kecelakaan pada variabel probable causes .....	128

<b>Gambar 4.11</b> Visualisasi grafik jumlah kecelakaan pesawat terbang komersial per tahun selama periode tahun 1918-2020.....	134
<b>Gambar 4.12</b> Visualisasi grafik jumlah kematian pesawat terbang komersial per tahun selama periode tahun 1918-2020.....	134
<b>Gambar 4.13</b> Visualisasi jumlah kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan scheduled revenue flight per tahun.....	137
<b>Gambar 4.14</b> Visualisasi jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan scheduled revenue flight per tahun.....	139
<b>Gambar 4.15</b> Visualisasi jumlah kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan charter / taxi flight per tahun.....	141
<b>Gambar 4.16</b> Visualisasi jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan charter / taxi flight per tahun.....	142
<b>Gambar 4.17</b> Visualisasi jumlah kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan postal flight per tahun.....	145
<b>Gambar 4.18</b> Visualisasi jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan postal flight per tahun.....	146
<b>Gambar 4.19</b> Visualisasi jumlah kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan cargo flight per tahun .....	148
<b>Gambar 4.20</b> Visualisasi jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan cargo flight per tahun .....	150
<b>Gambar 4.21</b> Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan technical factors .....	153

<b>Gambar 4.22</b> Grafik jumlah kematian pesawat terbang komersial per tahun kategori dengan penyebab kecelakaan technical factors.....	155
<b>Gambar 4.23</b> Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan conflict factors.....	157
<b>Gambar 4.24</b> Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan conflict factors.....	159
<b>Gambar 4.25</b> Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan human errors.....	161
<b>Gambar 4.26</b> Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan human errors.....	162
<b>Gambar 4.27</b> Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan poor weather.....	165
<b>Gambar 4.28</b> Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan poor weather.....	165
<b>Gambar 4.29</b> Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan disappeared without trace.....	167
<b>Gambar 4.30</b> Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan disappeared without trace.....	169
<b>Gambar 4.31</b> Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan crashing the other objects.....	171
<b>Gambar 4.32</b> Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan crashing the other objects.....	172

2

*This Page Intentionally Left Blank*

## LIST OF TABLES

<b>Tabel 4.1</b> Hasil perhitungan parameter pengukuran pemusatan data pada variabel years .....	105
<b>Tabel 4.2</b> Perhitungan parameter pemusatan data pada variabel total fatalities .....	108
<b>Tabel 4.3</b> Perhitungan parameter dispersi data pada variabel years.....	111
<b>Tabel 4.4</b> Perhitungan parameter dispersi data pada variabel total fatalities .....	114
<b>Tabel 4.5</b> Perhitungan parameter skewness dan kurtosis pada variabel years .....	117
<b>Tabel 4.6</b> Perhitungan parameter skewness dan kurtosis pada variabel total fatalities.....	119

*This Page Intentionally Left Blank*

2

# CHAPTER 1 INTRODUCTION



*This Page Intentionally Left Blank*



## CHAPTER 1

### INTRODUCTION

#### 1.1 Background of Study

Saat ini, pesawat terbang telah menjadi salah satu alat transportasi udara utama bagi masyarakat dunia. Selama beberapa dekade terakhir, industri penerbangan telah mengalami peningkatan volume dan frekuensi lalu lintas udara yang stabil di seluruh dunia, terkhusus operasi penerbangan komersial yang diperkirakan akan melipat gandakan volume dan kompleksitasnya selama periode tersebut. Menurut ICAO (*International Civil Aviation Organization*) [1], pada tahun 2019, terdapat sekitar 4.397 miliar penumpang di seluruh dunia yang menggunakan pesawat sebagai alat transportasi udara utama mereka. Data ini menunjukkan seberapa penting pesawat terbang sebagai salah satu transportasi udara bagi masyarakat dunia. Selain itu, diperkirakan bahwa permintaan perjalanan

udara akan terus mengalami peningkatan sekitar 1,8% per tahunnya [2].

Peningkatan permintaan tersebut terjadi bukan tanpa alasan. Hal ini dikarenakan transportasi udara disebut sebagai jenis transportasi paling aman dibandingkan jenis transportasi lainnya [3]. Menurut data Bureau of Transportation Statistics [4], selama periode 2004 – 2019, mode transportasi udara memiliki persentase kematian dan kecelakaan paling rendah dibandingkan mode transportasi lainnya. Hal ini dapat terjadi karena industri penerbangan diseluruh dunia telah melakukan peningkatan dan inovasi terhadap desain, keandalan, dan keamanan dari pesawat terbang yang digunakan [5]. Walaupun memiliki tingkat keamanan yang paling tinggi diantara mode transportasi lainnya, tetapi analisis terhadap kasus kecelakaan pesawat terbang sangat diperlukan untuk menghindari terjadinya kecelakaan pesawat terbang serupa dimasa yang akan datang.

Sejak dahulu, para penulis telah mencoba untuk menganalisis data kecelakaan pesawat terbang komersial yang pernah terjadi sebelumnya, dengan tujuan untuk dapat menemukan pola dan karakteristik dari data kecelakaan pesawat terbang komersial tersebut. Akan tetapi, analisis terhadap lokasi kecelakaan yang dilakukan hanya terbatas pada wilayah wilayah tertentu saja [6]. Hal ini dirasa kurang efektif, karena aktivitas

penerbangan pada saat ini tidak hanya dilakukan pada satu wilayah tertentu saja, melainkan dilakukan antar negara, bahkan benua sekalipun. Selain itu, metode analisis yang digunakan sebelumnya hanya menggunakan teknik pemakaran dan pengalaman saja, sehingga terkadang metode tersebut dapat mengarahkan pengguna data kepada kesalahan dalam melakukan penilaian terhadap hasil analisis data [7], serta metode tersebut juga belum bisa mengekstrak dan menemukan pola, karakteristik, dan wawasan yang tersembunyi pada data kecelakaan pesawat terbang komersial yang digunakan [8].

Untuk dapat menyelesaikan masalah-masalah tersebut, maka dibutuhkan metode analisis yang dapat menginvestigasi karakteristik dan mengekstrak wawasan yang tersembunyi dari data kecelakaan pesawat terbang komersial yang digunakan. Oleh karena itu metode analisis yang penulis digunakan adalah *descriptive statistics* [9] dan *exploratory data analysis (EDA)* [10]. Kedua metode analisis tersebut digunakan untuk mendeskripsikan data kecelakaan pesawat terbang komersial dengan bantuan perhitungan statistik, lalu mencari relasi dari masing-masing variabel yang digunakan. Dari masing-masing relasi tersebut, penulis akan melakukan analisis terhadap perubahan grafik jumlah kecelakaan dan kematian pesawat terbang komersial sebelum dan setelah dibentuknya ICAO.

Dari penulisan yang dilakukan, penulis membangun sebuah hipotesis awal, dimana grafik kecelakaan dan kematian pesawat terbang komersial pasti mengalami penurunan yang signifikan setelah dibentuknya ICAO sebagai organisasi keselamatan penerbangan sipil internasional.

## **1.2 Problem formulation**

Berdasarkan latar belakang penulisan mengenai kecelakaan pesawat terbang komersial sebelumnya, maka penulis dapat merumuskan masalah seperti berikut :

- Apa saja tahap-tahap yang dilakukan sebelum dapat menganalisis data kecelakaan pesawat terbang komersial yang digunakan ?
- Bagaimana bentuk karakteristik dan pola dari data kecelakaan pesawat terbang komersial yang digunakan ?
- Berapa banyak relasi antar variabel yang dapat dibuat dalam menganalisis data yang digunakan ?
- Apa saja yang dilakukan ICAO dalam menjalankan tugasnya sebagai organisasi keselamatan penerbangan sipil internasional ?
- Apakah terdapat penurunan terhadap grafik jumlah kecelakaan dan kematian pesawat terbang komersial,

serta intensitas kecelakaan pesawat terbang komersial setelah dibentuknya ICAO ?

### 1.3 Scope of the problems

- Raw data yang dikumpulkan berasal dari website BAAA (*Bureau of Aircraft Accidents Archives*).  
43
- Data yang digunakan untuk keperluan analisis adalah data kecelakaan pesawat terbang komersial dengan *region world*, yang dimulai dari tanggal 1 Januari 1918 – 31 Desember 2020.
- Metode analisis data yang digunakan adalah *descriptive statistics* dan *exploratory data analysis*.  
121
- Bahasa pemrograman yang digunakan adalah python, sementara *software* yang digunakan untuk menjalankan kode program adalah *jupyter notebook*.

### 1.4 The purposes of this study

- Menganalisis data kecelakaan pesawat terbang komersial BAAA yang sebelumnya hanya dikumpulkan saja.
- Menemukan wawasan dan fakta yang tersembunyi berdasarkan hasil analisis data kecelakaan pesawat terbang komersial BAAA.

- Mengimplementasikan penggunaan metode *descriptive statistics* dan *exploratory data analysis* dalam mendeskripsikan dan menganalisis data kecelakaan pesawat terbang komersial BAAA, sehingga dapat menjadi suatu informasi yang penting bagi pengguna data.
- Mengetahui karakteristik dari data kecelakaan pesawat terbang komersial BAAA yang digunakan. Menganalisis perubahan terhadap grafik jumlah kecelakaan dan kematian pesawat terbang komersial, serta intensitas kecelakaan sebelum dan setelah dibentuknya ICAO.

### **1.5 The benefits of this study**

- Menjadikan data BAAA menjadi data yang lebih informatif, karena dengan penulisan yang dilakukan akan dapat menghasilkan pengetahuan dan wawasan baru.
- Mengetahui hasil dari hipotesis yang dibangun terbukti benar atau salah
- Dapat mengimplementasikan dua metode analisis Descriptive Statistics dan Exploratory Data Analysis terhadap kasus kecelakaan pesawat terbang komersial diseluruh dunia dari data BAAA tahun 1918-2020

14

*This Page Intentionally Left Blank*



A front-facing view of a large commercial airplane, likely a Boeing 747, on a runway. The aircraft is centered in the frame, with its four engines and wings clearly visible. The sky is filled with heavy, grey clouds. In the foreground, there is a grassy area and a yellow sign with the letters 'W M W' and arrows pointing left and right. A red sign with '12R - 30L' is also visible near the runway.

## CHAPTER 2 LITERATURE REVIEW



*This Page Intentionally Left Blank*

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Related Studies

Dalam melakukan penyusunan penulisan, penulis membutuhkan inspirasi dan referensi dari penulisan-penulisan sebelumnya yang berkaitan dengan topik permasalahan pada penulisan yang sedang dilakukan. Penulisan sebelumnya yang berkaitan dengan topik permasalahan penulisan yang sedang dilakukan sekarang yaitu :

Penulisan yang dilakukan oleh Damien Kelly dan Marina Efthymiou pada tahun 2019, dengan judul penulisan “An analysis of human <sup>79</sup> factors in fifty controlled flight into terrain aviation accidents from 2007 to 2017” [5]. Penelitian ini

menentukan faktor-faktor yang terlibat dalam 50 kasus kecelakaan CFIT (Controlled Flight Into Terrain) dari 24 negara selama periode 10 tahun (2007-2017) <sup>124</sup> menggunakan metode HFACS (Human Factors Analysis and Classification). Hasil penulisan ini yaitu, salah satu faktor penting terjadinya kecelakaan CFIT ini adalah kondisi penglihatan dan pencahayaan, dimana ditemukan bahwa lebih dari 90% kecelakaan CFIT terjadi pada siang hari. Selain itu, pengalaman pilot juga merupakan elemen kunci dari kecelakaan ini, dimana standar jam terbang pilot berupa 5000 jam. Tetapi yang terjadi adalah 50% pilot memiliki jam terbang dibawah standar. Selanjutnya, menurut data analisis, 44% kecelakaan CFIT terjadi pada penerbangan pelayaran, dimana faktor manusia memiliki peran utama pada fase penerbangan ini. Hal ini dikarenakan pesawat perlu distabilkan pada penerbangan yang aman. Gangguan (distraction), angkuh pada kemampuan diri sendiri (complacency), dan kelelahan (fatigue) adalah elemen-elemen yang mungkin dirasakan oleh kru penerbangan sebagai

kontributor pada CFIT selama fase penerbangan ini. Penyebab lainnya dari CFIT ini adalah tidak terpasangnya perangkat peringatan terhadap medan (terrain) seperti <sup>47</sup>EGPWS (Enhanced Ground Proximity Warning System) dan GPWS (Ground Proximity Warning System) untuk mengurangi kecelakaan CFIT, terkhusus pada pesawat terbang general (general aviation aircraft) tidak memiliki perangkat peringatan tersebut. Hal ini dikarenakan tingginya harga yang harus dibayar untuk membeli dan menginstall perangkat tersebut, disebabkan bahwa kebanyakan pesawat jenis ini digunakan secara pribadi (privately owned and operated). Selain itu, tidak ada requirement secara internasional untuk mewajibkan perangkat tersebut terinstall pada jenis pesawat ini, tidak seperti penerbangan komersil yang mewajibkan perangkat ini terinstall pada pesawat mereka

Penulisan yang dilakukan oleh <sup>39</sup>Hyunseong Lee, Guoyi Li, Ashwin Rai dan Aditi Chattopadhyay pada tahun 2020 dengan judul penelitian “<sup>39</sup>Real-time anomaly detection framework using a

support vector regression for the safety monitoring of commercial aircraft” [11]. Penulisan ini bertujuan untuk memprediksi pola umum dalam performa penerbangan selama fase takeoff dan descent, yang mengizinkan diprediksinya anomali-anomali yang terjadi pada performa penerbangan. Hasil penulisan menggambarkan kekuatan model prediksi untuk mendeteksi anomali dan memberikan peringatan dini kepada awak pesawat selama penerbangan dimana perubahan mendadak dalam dinamika penerbangan dimungkinkan. Dimana model prediksi ini di compare dengan model optimisasi untuk mempelajari efisiensi dari model prediksi yang digunakan. Kekuatan model prediksi terlihat dimana tanggapan fitur pemantauan selama take off, dimana dinamika penerbangan berubah dengan cepat, seperti pada LGA yang diprediksi dan yang direkam, dimana LGA yang terekam berubah dari <sup>112</sup>-0,012 G menjadi 0,251 G dalam waktu 7 detik (nilai minimum dan maksimum penerbangan), dan model akan terlatih secara akurat memprediksi respons drastis ini 3 detik sebelumnya. Kemudian,

seleksi Fitur di gunakan berdasarkan pendekatan MICI dan berhasil mereduksi waktu komputational secara keseluruhan sebanyak 38.2%, dimana juga menambah akurasi prediksi (peningkatan sebesar 38% pada R2 dan penurunan pada RMSE sebesar 37%). Seleksi fitur tersebut kemudian digunakan untuk melatih model SVR dan memprediksi performa penerbangan. Waktu komputational dapat diekspansi karena metode MICI akan mengurangi keseluruhan fitur pelatihan dan melakukan penghapusan fitur-fitur yang tidak relevan. Model prediksi berhasil memprediksi skenario penerbangan yang tidak wajar dengan kerusakan mesin yang disimulasikan, dapat dideteksi dalam waktu 2 detik dengan fitur yang diprediksi. Dengan fitur yang diprediksi, anomali ditandai 1-3 detik sebelum kejadian sebenarnya dibandingkan dengan fitur yang direkam. Terdapat 4 parameter yang diuji menggunakan model prediksi yaitu LGA (Longitudinal Acceleration), VTA (Vertical Acceleration), LTA (Lateral Acceleration) dan PTCH (Pitch Angle). Model SVR juga akan dilatih dengan dua fitur yang berbeda, 47 fitur akan

dipilih melewati metode MICI, dan 98 fitur akan dipilih tanpa sebuah teknik fitur seleksi untuk mendemonstrasikan kapabilitas dari metode MICI. Model prediksi kemudian memprediksi LGA dan PTCH dengan lebih akurat dari VTA dan LTA, yang disebabkan oleh fluktuasi perilaku VTA dan LTA dibandingkan dengan LGA dan PTCH selama penerbangan.

Penulisan yang dilakukan oleh Kunimitsu Iwaware dan Tatsuo Oyama pada tahun 2015, dengan judul penulisan <sup>81</sup> “Statistical Data Analyses on Aircraft Accidents in Japan: Occurrences, Causes and Countermeasures” [12]. Penelitian ini bertujuan untuk menginvestigasi karakteristik-karakteristik utama dari kejadian, penyebab, dan penanggulangan kecelakaan pesawat terbang di Jepang. Penulisan ini menggunakan metode statistical data analysis dan mathematical modelling untuk menentukan hubungan antara perkembangan ekonomi, permintaan penerbangan, frekuensi kecelakaan pesawat atau helicopter, interval dari kecelakaan, dan jumlah kematian. Hasil dari penulisan berupa (1). Tindakan pencegahan kecelakaan

telah dikembangkan di ketiga jenis layanan lalu lintas udara yaitu, pesawat kecil, pesawat besar dan helicopter, yang sangat berkontribusi pada penurunan jumlah kecelakaan yang terjadi. Terkhusus pada kecelakaan pesawat kecil yang terjadi dikarenakan kesalahan pilot, dan kecelakaan pesawat besar selama takeoff dan landing. Selain itu ditemukan juga bahwa kecelakaan helicopter terjadi lebih sering selama fase penerbangan dibandingkan selama fase takeoff dan landing. (2). Kecelakaan yang disebabkan oleh factor cuaca tidak mengalami penurunan. Bahkan terjadi peningkatan pada jumlah dan persentase kecelakaannya. Hal ini mengindikasikan bahwa pengukuran untuk menentang factor cuaca sangat sulit untuk diimplementasikan, sehingga masih menjadi masalah utama yang harus diselesaikan. (3). Ditemukan bahwa penurunan jumlah kecelakaan dapat dicapai melalui berbagai upaya keselamatan yang dilakukan oleh setiap orang yang terlibat pada layanan lalu lintas udara, termasuk memperkuat aturan pada kebijakan public.inovasi teknologi pada pesawat, helicopter, dan



fasilitas sekitarnya seperti bandara dan radar pengatur udara, peningkatan metode Latihan untuk mekanik dan pengatur lalu lintas. (4). Setelah mengaplikasikan model analisis matematika, didapatkan bahwa pada kasus perkembangan ekonomi, permintaan pesawat dan jumlah kecelakaan, model yang digunakan dapat mengindikasikan bahwa adanya penurunan pada jumlah kecelakaan, bersamaan dengan perkembangan ekonomi dan permintaan penerbangan hingga tahun 1991. Selain itu ditemukan juga bahwa mengenai interval kecelakaan dan jumlah kematian, kecocokan yang baik untuk distribusi probabilitas diamati untuk masing-masing distribusi secara eksponensial dan binominal negative.

Penulisan yang dilakukan oleh <sup>115</sup> Tebogo Makaba, Wesley Doorsamy dan Babu Sena Paul pada tahun 2020, dengan judul penulisan <sup>75</sup> “Exploratory framework for analysing road traffic accident data with validation on Gauteng province data” [13]. Penulisan ini bertujuan untuk mengusulkan sebuah framework exploratory untuk menganalisis dataset kecelakaan lalu lintas

jalan menggunakan representasi grafik seperti boxplot, bar chart, dan histogram. Selain itu juga mengusulkan penggunaan <sup>87</sup> Principal Component Analysis (PCA) dan Linear Discriminant Analysis (LDA) untuk melakukan reduksi dimensi data (dimensionality reduction). Hasilnya adalah (1). PCA tidak memiliki kinerja yang lebih baik dibandingkan LDA, dimana LDA memiliki pengaruh yang baik terhadap dataset ketika menggunakan PCA untuk melakukan reduksi dimensi data. (2). Kemudian ketika menggunakan Naïve Bayes classifier, model classifier ini akan bekerja dengan baik pada data ketika Teknik dimensionality reduction LDA juga diterapkan pada data. (3). EDA dan dimensionality reduction dapat meningkatkan performa model yang diterapkan pada data

## 2.2 History of ICAO

Dahulu, setiap negara di seluruh dunia memiliki peraturan masing-masing mengenai siapa saja yang diperbolehkan untuk terbang di atas wilayah udara mereka. Hal tersebut tidak menjadi masalah ketika aktivitas penerbangan hanya dilakukan

didalam wilayah suatu negara tertentu saja. Tetapi hal tersebut akan menjadi masalah ketika aktivitas penerbagan dilakukan antar wilayah negara yang berbeda.

Selain itu, masalah lainnya adalah pada bidang keselamatan penerbangan. Dengan peningkatan sektor industri penerbangan yang sangat signifikan setiap tahunnya, keselamatan penerbangan masih menjadi perhatian penting bagi setiap indsutri penerbangan. Penerbangan internasional melibatkan setidaknya dua pihak, seperti pada penerbangan dari negara A ke negara B. Aktivitas penerbangan ini memerlukan pengawasan keselamatan yang tepat oleh kedua belah pihak. Fakta bahwa negara A memiliki tingkat keselamatan yang tinggi akan terasa sia-sia ketika otoritas keselamatan di negara B lalai dalam menjalankan tugasnya. Hal ini tidak hanya berdampak pada lalu lintas penerbangan di negara B sendiri, tetapi juga pada negara-negara lain ketika negara B merupakan tujuan penerbangan mereka. Oleh karena itu, diperlukan standar

regulasi dan aturan tertentu mengenai keselamatan penerbangan secara internasional yang perlu dipatuhi oleh setiap negara [14].

Untuk mengatasi masalah tersebut, dibentuk suatu organisasi yang dapat menstandarisasi aturan dan ketentuan pada sektor penerbangan secara internasional yang bernama ICAO (International Civil Aviation Organization) [15]. ICAO dibentuk pada tanggal 4 April 1947, melalui sebuah konvensi penerbangan sipil internasional yang dikenal sebagai Chicago Convention pada 7 Desember 1944, dan telah ditandatangani oleh 52 negara di dunia. Salah satu tujuan dari dibentuknya ICAO adalah untuk memastikan keselamatan dan pertumbuhan penerbangan sipil internasional yang teratur diseluruh dunia [1]. Regulasi, prosedur dan organisasi standar keselamatan yang dibuat oleh ICAO yang dimuat didalam SARPs (Standards and recommended practices), telah disetujui secara global oleh anggota-anggota ICAO, dan dimuat kedalam ICAO annexes [16].

### **2.3 Aviation Accident**

---

Menurut ICAO [17], dalam *Convention on International Civil Aviation* pada Annex 13, *standard of recommended practice*, kecelakaan penerbangan merupakan sebuah kejadian yang berhubungan dengan pengoperasian pesawat udara, dimulai ketika seseorang menaiki pesawat untuk tujuan terbang ke suatu tempat hingga ketika seluruh orang tersebut telah turun dari pesawat, dimana mesin atau baling-baling pesawat berhenti selama keadaan tersebut, yang menyebabkan :

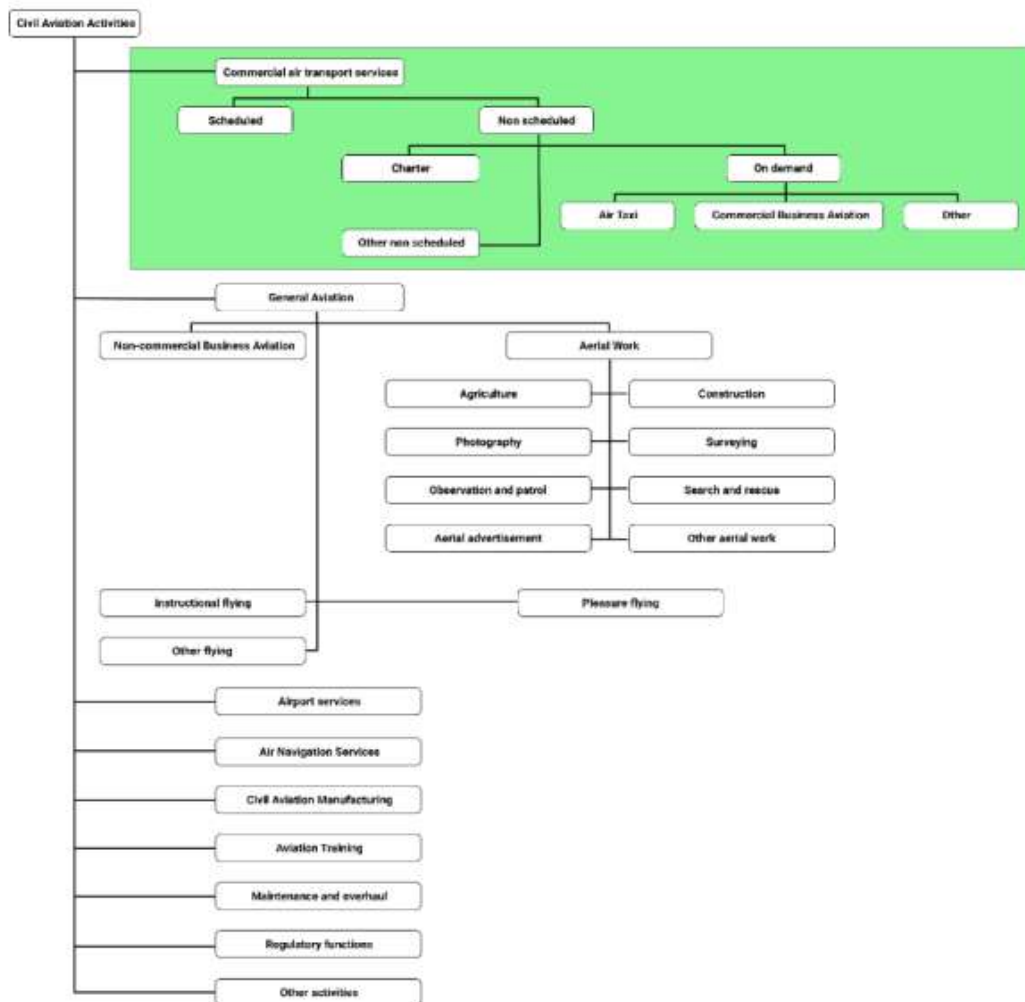
- Seseorang terluka <sup>82</sup> secara fatal atau serius sebagai akibat dari, berada didalam pesawat, kontak langsung dengan bagian manapun dari pesawat, dan paparan langsung terhadap ledakan jet. Terkecuali ketika seseorang tersebut terluka karena penyebab natural, disebabkan diri <sup>27</sup> sendiri atau ditimbulkan oleh orang lain, atau ketika cedera tersebut terjadi kepada penumpang gelap yang bersembunyi diluar area yang biasanya tersedia untuk penumpang dan awak pesawat, atau;

- Pesawat mengalami kerusakan atau kegagalan structural yang berdampak buruk pada kekuatan structural, performa atau karakteristik penerbangan dari pesawat, dan memerlukan perbaikan besar atau pergantian komponen pesawat yang terdampak. Terkecuali untuk kegagalan mesin atau kerusakan mesin, ketika kerusakan hanya terbatas pada satu mesin saja (termasuk penutup mesin dan aksesoris), untuk baling-baling, ujung sayap, antenna, ban, rem roda, panel, pintu pendaratan, kaca depan, dan lapisan pelindung pesawat, atau untuk kerusakan kecil pada bilah rotor utama, bilah rotor ekor, komponen pendaratan, dan yang dihasilkan dari hujan es atau serangan burung, atau;
- Pesawat hilang atau tidak dapat dilacak.

## 2.4 Civil Aviation

Penerbangan sipil merupakan jenis penerbangan yang mencakup penerbangan non militer [18]. ICAO membagi penerbangan sipil kedalam 9 jenis aktivitas penerbangan [19], dari kesembilan jenis aktivitas penerbangan sipil tersebut,

terdapat dua jenis aktivitas penerbangan yang membawa penumpang atau barang dengan jumlah banyak, yaitu penerbangan udara komersial (commercial aviation) dan penerbangan udara umum (general aviation). Adapun kesembilan jenis aktivitas penerbangan sipil tersebut dapat dilihat pada Gambar 2.1.



## **Gambar 2.1** Klasifikasi penerbangan sipil (civil aviation)

berdasarkan ICAO

### **2.4.1 Commercial Aviation**

Menurut ICAO pada Annex 6 [20], penerbangan komersial adalah pengoperasian sebuah pesawat udara pada satu atau lebih tahap secara terjadwal atau tidak terjadwal, yang tersedia untuk publik dengan imbalan dan untuk disewa yang melibatkan pengangkutan penumpang, cargo atau surat. Operasi pesawat terbang komersial dapat diklasifikasikan kedalam dua jenis, yaitu operasi terjadwal (scheduled) dan tidak terjadwal (non scheduled).

Penerbangan terjadwal adalah layanan penerbangan udara yang dilakukan untuk mendapatkan imbalan sesuai dengan jadwal yang telah diterbitkan, atau merupakan penerbangan yang teratur atau sering, sehingga merupakan rangkaian sistematis yang dapat dikenali, dan terbuka untuk pemesanan langsung oleh anggota masyarakat [19].



Sementara penerbangan tidak terjadwal merupakan layanan penerbangan udara komersial yang dilakukan secara berbeda dari layanan penerbangan udara komersial yang terjadwal. Penerbangan non scheduled terbagi menjadi charter dan on demand. Penerbangan charter adalah operasi penerbangan tak terjadwal yang menggunakan pesawat sewaan. Sementara penerbangan on demand merupakan penerbangan tak terjadwal yang digunakan untuk pengangkutan penumpang, barang, atau surat yang dilakukan dengan pesawat yang lebih kecil termasuk helicopter [19].

#### **2.4.2 General Aviation**

Penerbangan umum didefinisikan sebagai penerbangan untuk tujuan statistik, bukan untuk mendapatkan imbalan atau disewa. Berdasarkan tujuannya, aktivitas penerbangan umum dibagi menjadi *instructional flight*, *business flight*, *pleasure flight*, *aerial work* dan penerbangan lainnya [19].

*Instructional flight* didefinisikan sebagai penggunaan sebuah pesawat terbang untuk tujuan instruksi penerbangan formal dengan seorang instruktur, yang biasa dilakukan oleh *aero-clubs* dan sekolah penerbangan. Kemudian *pleasure flying* didefinisikan sebagai penggunaan sebuah pesawat terbang untuk keperluan personal atau rekreasi yang tidak memiliki hubungan dengan profesi atau bisnis. Sementara *business flying* didefinisikan sebagai penggunaan sebuah pesawat terbang untuk membawa personil dan properti untuk memenuhi kebutuhan transportasi pejabat suatu bisnis, firma, perusahaan atau korporasi.

## **2.5 Data Analysis**

*Data analysis* atau analisis data merupakan sebuah proses untuk memeriksa, membersihkan, mentransformasi, dan memodelkan suatu data untuk dapat menemukan informasi yang berguna, membuat kesimpulan, dan mendukung pengambilan keputusan selanjutnya untuk menyelesaikan permasalahan yang ada [21]. Analisis data telah menjadi dasar dalam melakukan

investigasi terhadap suatu data di berbagai bidang pengetahuan. Proses ini juga biasa digunakan pada bidang bisnis, dimana perlunya melakukan analisis data bisnis secara rutin untuk mendukung pengambilan keputusan agar bisnis yang dikelola dapat berkembang [22]. Terdapat 5 tahap dalam melakukan analisis data, yaitu *determine analysis objectives*, *data collection*, *data preprocessing*, *analyze the data*, dan *interpret the results*.

### **2.5.1 Determine Analysis Objectives**

Tahap pertama untuk melakukan analisis data adalah menentukan tujuan dari analisis data, lalu membuat hipotesis sementara yang akan diuji hasilnya dengan metode analisis data yang akan digunakan. Dengan menentukan tujuan melakukan analisis data, akan mempermudah penulis untuk memilih jenis data, membersihkan dan mentransformasikan data, dan menganalisis <sup>99</sup> data, yang akan digunakan untuk menguji

hipotesis yang telah dibangun, serta membuat kesimpulan akhir dari data [23].

### 2.5.2 Data Collection

Data collection merupakan sebuah proses mengumpulkan dan mengukur informasi mengenai variabel yang akan digunakan untuk menganalisis sebuah data, yang didapatkan dari berbagai sumber yang berbeda. Proses ini memungkinkan penulis untuk dapat menjawab pertanyaan seputar penulisan yang dilakukan, menguji hipotesis, menentukan pendekatan analisis yang akan diterapkan, serta dapat mengevaluasi hasil dan membuat prediksi trend dimasa depan [24].

Dalam mengumpulkan informasi, sebuah data dapat dikategorikan data kedalam dua kategori berbeda, yaitu data *primary* dan *secondary*. Data *primary* atau primer merupakan kategori data yang dikumpulkan pertama kali secara langsung

oleh seseorang dengan tujuan untuk mendapatkan solusi dari masalah yang dihadapi, sementara data *secondary* atau <sup>32</sup> *sekunder* merupakan data yang telah dikumpulkan oleh orang lain untuk tujuan tertentu, seperti menganalisis statistik resmi dari suatu organisasi, melihat rekaman administrasi, dan tugas lainnya [25]. Untuk dapat mengumpulkan data primer, dapat menggunakan metode seperti survey, observasi, kuisisioner, atau *interview* secara personal. Sementara untuk data sekunder, dapat dikumpulkan melalui *website*, buku, artikel, atau jurnal [26].

### 2.5.3 Data Preprocessing

Data yang telah dikumpulkan sebelumnya, perlu melewati tahap *data preprocessing* sebelum dapat dianalisis lebih lanjut. Hal ini dikarenakan kebanyakan *real world data* merupakan data yang tidak terstruktur dan sulit diinterpretasikan tanpa bantuan data preprocessing [27]. Sehingga *data preprocessing* merupakan proses yang penting dilakukan sebelum memulai proses analisis data [22].

*Data preprocessing* memiliki tujuan untuk menyelesaikan masalah yang terjadi pada data mentah seperti data yang tidak konsisten, *noise*, data yang hilang, dimensi terlalu tinggi dan data yang tidakimbang [28]. Secara umum, *data preprocessing* terbagi kedalam 4 teknik, yaitu *data cleaning*, *data reduction*, *data transformation* dan *data integration* [29]. Penggunaan keempat bagian tersebut bergantung terhadap kondisi data yang dimiliki.

### **2.5.3.1 Data Cleaning**

*Data cleaning* merupakan sebuah teknik untuk menemukan dan memperbaiki data yang tidak lengkap, hilang, data yang berulang, serta outliers. Tujuan dilakukannya *data cleaning* untuk mereduksi waktu dalam mengeksekusi kode program pada data, mereduksi kompleksitas data, dan meningkatkan kualitas dari data yang dimiliki [27].

### **2.5.3.2 Data Reduction**

*Data reduction* atau reduksi data merupakan sebuah teknik untuk menampilkan data yang dimiliki dalam dimensi yang lebih sedikit dari data aslinya, tetapi tetap dapat menghasilkan hasil analisis yang sama. Proses data reduction dapat dilakukan dengan cara mengeliminasi fitur atau atribut yang memiliki pengaruh sangat sedikit atau tidak sama sekali terhadap analisis data [27]. Beberapa teknik yang dapat digunakan untuk mereduksi data adalah *feature selection*, *feature extraction* dan *dimensionality reduction* [30].

### **2.5.3.3 Data Transformation**

*Data transformation* atau transformasi data merupakan sebuah teknik untuk mengubah nilai dan jenis data pada data yang digunakan kedalam format yang sesuai dengan kebutuhan analisis data. Teknik ini dilakukan agar data yang digunakan cocok untuk dikombinasikan dengan jenis data lainnya, sehingga didapatkan informasi dan wawasan baru yang tersembunyi pada data. [27]. Teknik yang dapat digunakan

untuk mentransformasi data adalah *normalization*, *smoothing*, dan juga *encoding* [28].

#### **2.5.3.4 Data Integration**

*Data integration* merupakan sebuah teknik untuk menjaga integritas data. Hal ini dikarenakan data dikumpulkan dari berbagai sumber, maka dapat menyebabkan adanya data yang redundan dan tidak konsisten, dimana dapat memperburuk akurasi analisis data. Oleh karena itu, data dari berbagai sumber tersebut akan dikombinasikan menjadi sebuah dataset yang koheren dan terintegrasi dengan baik [28].

#### **2.5.4 Descriptive Statistics**

Sebelum membahas mengenai *descriptive statistics*, penulis perlu mengetahui apa itu *statistics*. *Statistics* atau statistika merupakan ilmu mengumpulkan, menganalisis,



menyajikan dan menafsirkan data [31]. Selanjutnya, *statistics* dibagi menjadi beberapa kategori, salah satunya adalah *descriptive statistics*. *Descriptive statistics* merupakan sebuah proses mendeskripsikan data, dengan merangkum dan menyederhanakan masing-masing variabel pada data yang digunakan sedemikian rupa, sehingga dapat disajikan menjadi suatu informasi yang lebih menarik, bermanfaat, dan mudah dimengerti oleh pengguna data [32]. Proses deskripsi data yang digunakan dapat disajikan kepada pengguna data melalui perhitungan numerik, grafik atau tabel. Ketika melakukan analisis suatu data tanpa melalui tahap *descriptive statistics* dapat menyebabkan hilangnya banyak wawasan yang berharga dari suatu data tersebut, sehingga seringkali mengarahkan pengguna data pada kesimpulan yang keliru [33].

Terdapat beberapa teknik untuk dapat mendeskripsikan suatu data. Tetapi sebelum mengaplikasikan teknik-teknik tersebut pada data yang akan dianalisis, penulis harus mengetahui terlebih dahulu jenis data pada setiap variabel yang

dimiliki. Suatu data secara umum terbagi kedalam dua jenis, yaitu data *quantitative* dan *qualitative*. Data *quantitative* merupakan jenis data dimana variabel didalamnya direpresentasikan dalam bentuk numerik atau angka. Sedangkan data *qualitative* merupakan jenis data dimana variabel yang berada didalamnya direpresentasikan melalui pengelompokan secara verbal [34]. Data *qualitative* dibagi kedalam dua jenis variabel, yaitu nominal dan ordinal. Variabel nominal memiliki dua atau lebih kategori yang ditulis dengan huruf dan tidak memiliki urutan tertentu. Sementara data ordinal mirip dengan data nominal, tetapi dapat diurutkan dengan ketentuan tertentu [9].

#### **2.5.4.1 Describing Quantitive Data**

Secara umum, terdapat beberapa jenis pengukuran untuk mendeskripsikan data *quantitative*, yaitu *measures of frequency*, <sup>60</sup> *measures of central tendency, measures of dispersion, skewness*, dan *kurtosis* [35]. Masing-masing pengukuran tersebut memiliki

tujuan yang sama untuk mendeskripsikan suatu data, dan menjadikan data tersebut sebagai informasi yang penting bagi pengguna data.

#### **2.5.4.1.1 Measures of Frequency**

*Measures of frequency* merupakan salah satu teknik pengukuran statistik deskriptif dengan membuat daftar setiap nilai atau kategori berbeda pada suatu variabel, lalu menghitung frekuensi kemunculannya pada data yang digunakan [36]. Pengukuran frekuensi atau jumlah kemunculan digunakan untuk mendeskripsikan sebuah variabel tunggal [37], yang terdiri dari informasi personal variabel tersebut.

94

#### **2.5.4.1.2 Measures of Central Tendency**

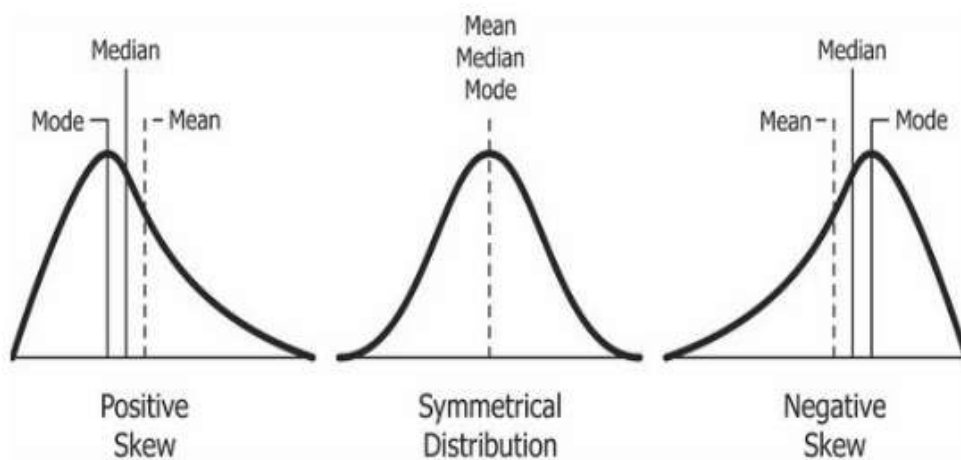
*Measures of central tendency* merupakan teknik pengukuran statistik untuk menentukan lokasi titik tengah atau titik pusat pada suatu distribusi data. Terdapat tiga jenis parameter pengukuran titik pusat dari suatu distribusi data, yaitu

*mean*, *median*, dan *mode* [38]. Untuk dapat menentukan parameter pengukuran yang tepat untuk menentukan titik pusat dari suatu distribusi data, bergantung pada jenis distribusi data tersebut. Distribusi data terbagi menjadi 2 jenis, yaitu *Symmetrical distribution* dan *skewed*.

*Symmetrical distribution* merupakan distribusi data normal dan simetris, yang ditandai dengan nilai *mean*, *median* dan *mode* yang sama. Sementara jenis distribusi data *skewed* merupakan distribusi data yang miring dan tidak simetris, yang terbagi menjadi dua jenis, yaitu *positive skewed* dan *negative skewed*. *Positive skewed* ditandai dengan nilai *mean* lebih besar dari nilai *median*, serta nilai *median* lebih besar dari nilai *mode*. Sementara *negative skewed* ditandai dengan nilai *mean* lebih kecil dari nilai *median*, dan nilai *median* lebih kecil dari nilai *mode*. Ilustrasi ketiga jenis distribusi data tersebut dapat dilihat pad **Gambar 2.2**.

Suatu variabel yang memiliki jenis distribusi data *symmetrical distribution* dapat menggunakan nilai dari salah

satu parameter *mean*, *median* dan *mode*. Sedangkan variabel yang memiliki jenis distribusi data skewed dapat menggunakan nilai dari parameter median. Hal ini dikarenakan jenis distribusi data ini memiliki outliers, dimana hanya parameter median yang tidak memiliki pengaruh yang besar terhadap hadirnya outliers [39].



**Gambar 2.2** Jenis distribusi suatu data

*Mean* adalah salah satu parameter pengukuran yang digunakan untuk menentukan titik pusat suatu distribusi data. *Mean* dapat dihitung dengan cara menjumlahkan seluruh nilai

pada data, lalu membaginya dengan jumlah sampel nilai [39].

Rumus perhitungan *mean* <sup>20</sup> dapat dilihat pada persamaan (2.1).

$$\bar{X} = \frac{\sum X}{N} \quad (2.1)$$

Kelebihan dari parameter *mean* adalah karena parameter ini menggunakan seluruh nilai didalam data, sehingga baik untuk merepresentasikan suatu data. Sementara kelemahan dari parameter *mean* adalah sangat sensitif terhadap nilai ekstrim atau *outliers*, sehingga tidak cocok digunakan pada variabel yang memiliki jenis distribusi data *skewed* [40].

Selanjutnya, median merupakan sebuah nilai yang berada di posisi tengah ketika seluruh nilai sampel disusun, baik secara *ascending* atau *descending* [39]. Sebelum menghitung nilai median, nilai pada setiap sampel data perlu diurutkan terlebih dahulu, baru kemudian dapat menghitung nilai median. Dalam menghitung nilai median, terdapat dua ketentuan, yaitu ketika sampel data berjumlah genap dan ganjil. Rumus

menghitung median untuk sampel data yang berjumlah genap terlihat pada persamaan (2.2), sementara pada sampel data yang berjumlah ganjil terlihat pada persamaan (2,3).

$$\text{Median} = \left(\frac{n}{2}\right)^{\text{th}} \text{ observation} \quad (2.2)$$

$$\text{Median} = \frac{\left(\frac{n}{2}\right)^{\text{th}} + \left(\frac{n+1}{2}\right)^{\text{th}} \text{ observation}}{2} \quad (2.3)$$

Sama seperti parameter *mean*, parameter *median* juga memiliki kelebihan dan kekurangan. Kelebihan dari parameter *median* adalah mudah dihitung, tidak terdistorsi oleh nilai ekstrim atau outliers, dan dapat digunakan baik pada jenis data *quantitative* maupun *qualitative*. Kelemahannya, parameter *median* tidak menggunakan seluruh nilai sampel yang berada pada data, sehingga tidak seluruh informasi pada data digunakan secara maksimal [41].



Parameter pengukuran titik pusat pada distribusi data yang terakhir adalah *mode*. *Mode* didefinisikan sebagai nilai yang memiliki frekuensi kemunculan paling sering didalam data. Rumus untuk menghitung mode terlihat pada persamaan (2,4)

$$Mode = l + h \left( \frac{f_m - f_1}{2f_m - f_1 - f_2} \right) \quad (2.4)$$

Kelebihan dari parameter *mode* adalah hanya nilai *mode* yang dapat digunakan untuk menentukan titik pusat pada jenis data *qualitative* nominal, serta dapat dihitung dengan mudah [41].

#### 2.5.4.1.3 Measures of Dispersion

Pengukuran central tendency atau titik tengah dari sebuah distribusi data sebelumnya belum cukup untuk dapat mendeskripsikan keseluruhan dari suatu variabel didalam suatu data. Hal ini dikarenakan dua buah data dapat memiliki rata-rata

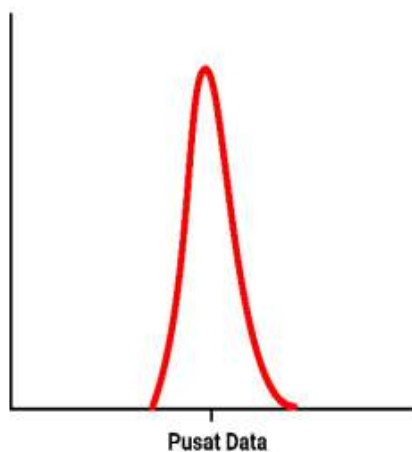


yang sama, tetapi sebenarnya keduanya sangat berbeda [42]. Sehingga untuk mendeskripsikan data dengan baik, perlu diketahui tingkat disperse atau persebaran dari setiap nilai pada data tersebut. Oleh karena itulah digunakan perhitungan untuk menentukan tingkat dispersi atau persebaran data dari titik pusat distribusi datanya dengan *measures of dispersion* [43].

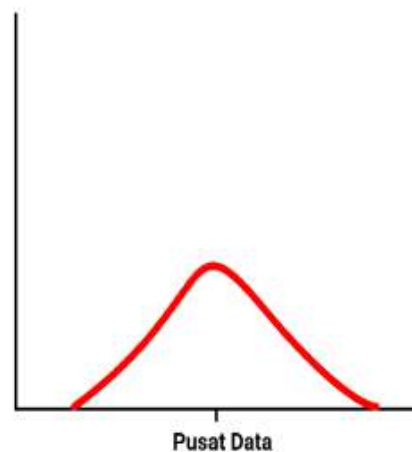
Dalam ilmu statistika, *dispersion*, *spread*, atau *variability* memiliki sinonim dengan lebar dari distribusi suatu data. Ketika distribusi suatu data memiliki dispersi yang lebih rendah, maka setiap nilai pada data tersebut lebih konsisten dan cenderung mengelompok disekitar titik pusat distribusi data. Sementara ketika dispersi pada sebuah distribusi data lebih tinggi, maka setiap nilai pada data tersebut akan tidak konsisten, dan meningkatkan adanya nilai ekstrim atau outliers. Sehingga distribusi data cenderung menyebar lebih jauh dari titik pusat distribusi data seperti yang tampak pada **Gambar 2.3**.

Terdapat beberapa jenis parameter pengukuran persebaran data seperti range, interquartiles range (IQR),

variance, dan standard deviation [44]. Penggunaan masing-masing parameter pengukuran dispersi tersebut bergantung dari jenis distribusi data. Untuk variabel yang memiliki jenis distribusi data simetris dapat menggunakan variance atau standard deviation untuk menentukan persebaran data. Sementara variabel yang memiliki jenis distribusi data skewed dapat menggunakan parameter IQR untuk menentukan persebaran data. Hal ini dikarenakan IQR tidak memiliki pengaruh yang besar terhadap outliers [44].



A



B

**Gambar 2.3** Persebaran setiap nilai dari titik pusat distribusi data pada kondisi (a). Low dispersion; (b). High dispersion

*Range* merupakan jarak antara nilai observasi tertinggi dan terendah pada data. Semakin tinggi nilai *range*, menunjukkan semakin besar jarak antara nilai tertinggi dan terendah pada data [45]. Rumus untuk menghitung nilai *range* terlihat pada persamaan (2.5)

$$\text{Range}(X) = \text{Max}(X) - \text{Min}(X) \quad (2.5)$$

Kelebihan penggunaan *range* sebagai parameter pengukuran penyebaran data adalah mudah dihitung. Tetapi *range* memiliki beberapa kelemahan, diantaranya parameter *range* sangat sensitif terhadap adanya outliers, selain itu parameter *range* tidak menggunakan seluruh nilai observasi pada data yang digunakan [43].

Parameter selanjutnya adalah IQR. IQR merupakan salah satu parameter pengukuran penyebaran data berdasarkan kuartil pertama dan kuartil ketiga pada data, dimana IQR merupakan selisih diantara kedua kuartil tersebut [46]. Rumus untuk menghitung nilai IQR terlihat pada persamaan (2.6).

$$IQR = Q_3 - Q_1 \quad (2.6)$$

Karena berada diantara Q1 dan Q3, maka dengan kata lain, IQR mencakup 50% data yang berada diantara Q1 dan Q3. Parameter pengukuran IQR mirip dengan *median*, dimana keduanya tidak terpengaruh oleh adanya nilai ekstrim atau *outliers*. Sehingga ketika memiliki variabel dengan jenis distribusi data *skewed*, maka IQR merupakan parameter pengukuran yang tepat untuk menentukan persebaran atau dispersi data [44].

Parameter *variance* dihitung berdasarkan kuadrat dari total dari setiap data yang dikurangi dengan nilai *mean*,

kemudian dibagi dengan jumlah sampel data yang dimiliki [45]. Pengkuadratan dilakukan karena jika hanya mengurangi setiap nilai data dengan *mean*, lalu ketika seluruh hasilnya dijumlahkan maka akan menghasilkan nilai negatif, sehingga untuk mengatasi hal tersebut, hasil pengurangan akan dikuadratkan agar hasilnya tetap positif [47]. Terdapat dua rumus untuk menghitung parameter *variance*, tergantung apakah penulis akan menghitung *variance* untuk seluruh populasi data, atau menggunakan sampel untuk memperkirakan *variance* dari populasi data.

Untuk menghitung *variance* untuk seluruh populasi data, maka hasil pengurangan nilai setiap data dengan nilai *mean* akan dikuadratkan, lalu dibagi dengan seluruh jumlah sampel data, dengan rumus seperti pada persamaan (2.7).

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad (2.7)$$

Sementara untuk menghitung nilai sampel untuk memperkirakan *variance* dari populasi data, rumus yang digunakan sama seperti rumus untuk menghitung *variance* untuk seluruh populasi data, tetapi dibagi dengan seluruh jumlah sampel data dikurangi dengan 1. Hal ini dilakukan karena data sampel tidak memiliki jumlah yang pasti dibandingkan populasi data [44]. Untuk menghitung nilai sampel *variance*, dapat menggunakan rumus seperti pada persamaan (2.8).

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (2.8)$$

*Standard deviation* merupakan salah satu parameter pengukuran penyebaran data yang digunakan untuk menyelesaikan permasalahan yang dimiliki oleh parameter pengukuran *variance*. Untuk menghitung nilai *variance*, digunakan pengkuadratan yang menyebabkan hasil perhitungan *variance* tidak memiliki satuan pengukuran yang sama dengan



satuan pengukuran data aslinya. Sehingga digunakanlah parameter pengukuran lain yaitu *standard deviation* untuk menyelesaikan masalah *variance*. *Standard deviation* akan menggunakan akar untuk mengembalikan satuan *variance* yang kuadrat menjadi satuan data aslinya [33].

*Standard deviation* menggambarkan seberapa berbeda jarak antara setiap data yang dimiliki terhadap nilai *mean* [48]. Semakin rendah nilai *standard deviation*, maka jarak antara setiap nilai yang ada pada data yang dikelompokkan lebih dekat. Sedangkan semakin tinggi nilai *standard deviation*, maka setiap nilai pada data akan lebih menyebar karena jarak standar yang dimiliki juga lebih besar [44].

Kelebihan penggunaan *standard deviation* adalah dapat mendeteksi dan menghitung nilai *skewness* dengan bantuan nilai *mean*. Sementara kelemahan *standard deviation* adalah tidak cocok digunakan untuk menghitung penyebaran data pada variabel dengan jenis distribusi data *skewed*. Hal ini dikarenakan nilai *mean* akan sensitif jika terdapat *outliers* [43]. Untuk dapat

menghitung nilai *standard deviation*, digunakan akar kuadrat dari nilai *variance* yang telah didapatkan, dengan rumus seperti pada persamaan (2.9).

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (2.9)$$

#### 2.5.4.1.4 Skewness dan Kurtosis

*Skewness* dan *kurtosis* merupakan salah satu parameter pengukuran untuk mendeskripsikan karakteristik bentuk dari sebuah distribusi data [49]. *Skewness* digunakan untuk mengukur seberapa miring distribusi data yang dimiliki dibandingkan distribusi data normal, baik miring ke sisi kanan maupun ke sisi kiri. Dengan kata lain, *skewness* menunjukkan karakteristik kepadatan dari kurva distribusi data yang dimiliki [50].

*Skewness* dapat memiliki nilai nol, positif, dan negatif. Ketika *skewness* memiliki nilai nol, maka jenis distribusi data



yang dimiliki adalah simetris. Tetapi ketika nilai *skewness* tidak nol, maka terdapat beberapa aturan mengenai *range* nilai *skewness* sebagai berikut

- Ketika nilai *skewness* kurang dari -1 atau lebih dari 1, maka distribusi data sangat miring (*highly skewed*);
- Ketika nilai *skewness* berada diantara -1 dan -0.5, atau diantara 0.5 dan 1, maka distribusi data cukup miring (*moderately skewed*);
- Ketika nilai *skewness* berada diantara -0.5 dan 0.5, maka distribusi data kurang lebih simetris; [51]

Dari *range* nilai *skewness* yang dijelaskan sebelumnya, ketika *skewness* memiliki nilai positif atau berada di *range* positif, maka kurva distribusi data akan miring ke kanan (*positive skewed*). Sementara ketika *skewness* memiliki nilai negatif atau berada di *range* negatif, maka kurva distribusi data akan miring ke kiri (*negative skewed*) [52]. Ketiga karakteristik kurva distribusi data tersebut dapat dilihat pada **Gambar 2.2**.

Kemudian parameter *kurtosis* digunakan untuk menunjukkan puncak dan ketebalan ekor kurva distribusi data [50]. Nilai kurtosis dapat menentukan apakah suatu distribusi data memiliki ekor kurva yang tebal (*heavy-tailed*) atau tipis (*light-tailed*). Sebuah variabel dengan ekor kurva distribusi yang tebal memiliki nilai ekstrim atau *outliers* dengan jumlah yang lebih banyak. Sedangkan variabel dengan ekor kurva distribusi yang tipis memiliki nilai ekstrim atau *outliers* dengan jumlah yang lebih sedikit.

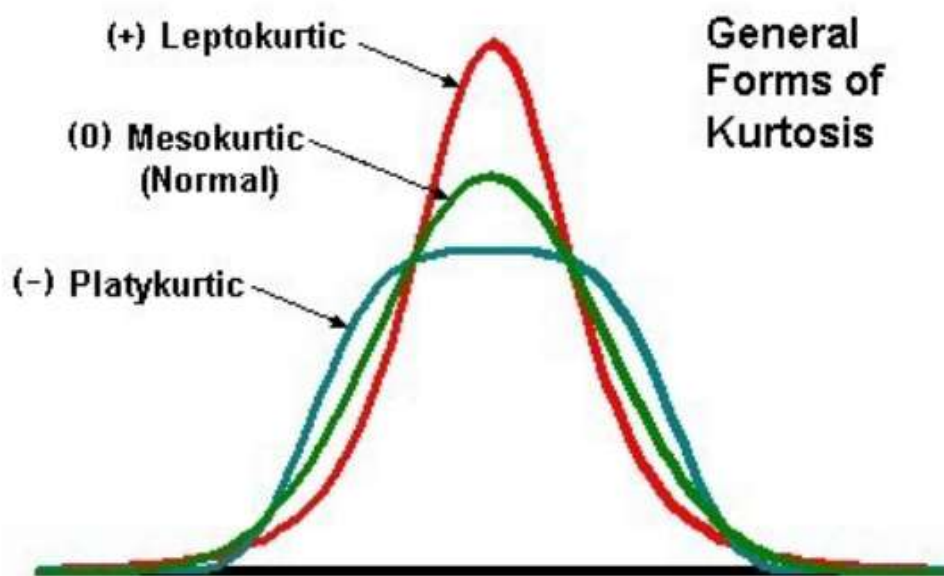
Terdapat tiga jenis distribusi data berdasarkan parameter kurtosis, yaitu *mesokurtic*, *platykurtic* dan *leptokurtic*. *Mesokurtic* merupakan jenis distribusi data normal yang memiliki nilai kurtosis 3. *Platykurtic* merupakan jenis distribusi data yang memiliki nilai kurtosis negatif atau kurang dari 3, yang diindikasikan dengan ekor kurva yang tipis dan memiliki *outliers* dengan jumlah yang sedikit. Sedangkan *leptokurtic* merupakan jenis distribusi data yang <sup>32</sup>memiliki nilai kurtosis lebih besar dari 3, diindikasikan dengan ekor kurva yang tebal,

serta memiliki outliers dengan jumlah yang lebih banyak [53].

Ketiga jenis distribusi data tersebut dapat dilihat pada Gambar

#### 2.4.

Dari penjelasan sebelumnya, dapat diketahui bahwa perbedaan utama dari *kurtosis* dan *skewness* adalah *skewness* mengacu kepada derajat simetri dari kurva distribusi data, sedangkan *kurtosis* lebih mengacu kepada derajat kehadiran *outliers* berdasarkan kurva distribusi data yang digunakan [54].



Gambar 2.4 Distribusi data berdasarkan nilai kurtosis

#### 2.5.4.2 Describing Qualitative Data

Jenis data selanjutnya adalah data *qualitative* atau data *categorical*. Karena jenis data *categorical* terdiri dari nilai yang direpresentasikan tidak dalam bentuk numerik [34], maka tidak masuk akal untuk menghitung parameter pengukuran pemusatan data seperti *mean*, pengukuran penyebaran data seperti *range*, IQR *standard deviation*, dan *variance*, serta pengukuran kurva distribusi data seperti *skewness* dan *kurtosis* [9]. Penulis akan mendeskripsikan variabel yang memiliki jenis data *qualitative* atau *categorical* dengan menggunakan frekuensi kemunculan, persentase, dan proporsi dari masing-masing kategori pada data [36].

### **2.5.5 Exploratory Data Analysis (EDA)**

Metode exploratory data analysis (EDA) Sebagian besar dikembangkan dari seorang ahli statistika terkenal bernama John Tukey dan rekannya, yang disajikan dalam sebuah karya klasik mereka yang berjudul “Exploratory Data Analysis” [55]. Metode ini memiliki tujuan untuk memahami struktur dan pola

dari data, dan melayani tujuan analitik seperti mengidentifikasi outliers, menemukan penyimpangan pada data, menentukan variabel yang penting, menyarankan hipotesis awal yang akan dibangun, mengidentifikasi transformasi yang perlu dilakukan pada data, dan menyarankan model yang sesuai untuk diterapkan pada suatu data [56]. Penggunaan EDA dapat direpresentasikan dalam bentuk analisis statistik atau visualisasi grafik untuk menginvestigasi pola yang tersembunyi pada data [13], menjawab hipotesis dan asumsi-asumsi yang dibangun, ataupun mengungkapkan trend pada data yang digunakan [57].

14

*This Page Intentionally Left Blank*



**CHAPTER 3**  
**RESEARCH METHODOLOGY**



*This Page Intentionally Left Blank*



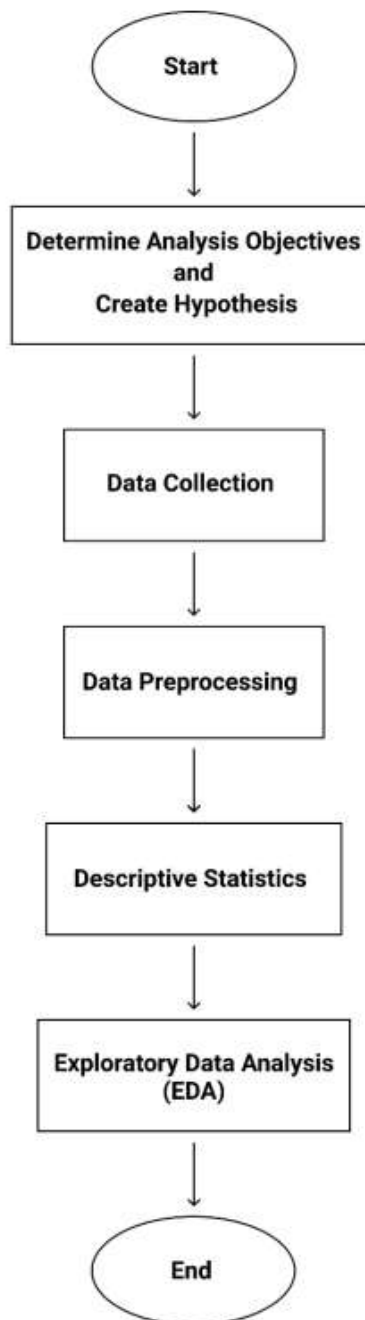
## CHAPTER 3

### RESEARCH METHODOLOGY

Pada penulisan yang sedang dilakukan, terdapat beberapa tahapan yang dilakukan oleh penulis dalam menganalisis data kecelakaan pesawat terbang komersial BAAA. Tahapan-tahapan penulisan tersebut ditunjukkan dalam bentuk *flowchart* seperti pada **Gambar 3.1**.

Proses-proses tersebut terdiri atas *determine analysis objectives and create hypothesis, data collection, data preprocessing, descriptive statistics, dan exploratory data analysis*. Seluruh tahapan penulisan tersebut diaplikasikan

menggunakan bahasa pemrograman python. Python merupakan salah satu bahasa pemrograman tingkat tinggi yang *interpreted*, beorientasi pada objek, dan memiliki *semantics* yang dinamis [58]. Python dipilih karena memiliki *syntax* yang lebih sederhana dibandingkan bahasa pemrograman lainnya



**Gambar 3.1** Tahap Penulisan

Dalam menggunakan bahasa pemrograman python, terdapat banyak *library* yang telah disediakan untuk

menjalankan fungsi tertentu. Pada penulisan ini, penulis hanya menggunakan beberapa *library* saja seperti *pandas*, *NumPy*, *seaborn* dan *matplotlib*. *Pandas* yang merupakan singkatan dari panel dan data, merupakan sebuah *library* yang digunakan untuk menganalisis dan mengkalkulasikan statistik dari data multi dimensi [59]. Selain *pandas*, terdapat salah satu *library* yang digunakan untuk komputasi data, yaitu adalah *NumPy* (*Numerical Python*) yang berfokus pada penyimpanan dan pengolahan data dalam berbentuk *array* [60]. Selanjutnya terdapat juga *library* yang digunakan untuk melakukan visualisasi data, yang dinamakan *matplotlib* dan *seaborn* [61].

### **3.1 Determine Analysis Objectives and Create Hypothesis**

Tahap pertama yang akan dilakukan pada analisis data adalah menentukan tujuan dari analisis data, lalu membangun hipotesis awal dari data tersebut. Tujuan penulis melakukan analisis data kecelakaan pesawat terbang komersial BAAA adalah dapat mendeskripsikan karakteristik dari data kecelakaan

pesawat terbang komersial BAAA dengan bantuan perhitungan statistik. Lalu mencari relasi setiap variabel pada data untuk menganalisis perubahan grafik jumlah kecelakaan dan kematian pesawat terbang komersial sebelum dan setelah dibentuknya ICAO sebagai badan keselamatan penerbangan sipil internasional.

Kemudian hipotesis awal yang dibangun oleh penulis adalah, tentunya terjadi penurunan grafik jumlah kecelakaan dan kematian pesawat terbang komersial setelah dibentuknya ICAO. Hal ini dikarenakan ICAO telah menstandarisasi aturan dan ketentuan mengenai penerbangan sipil internasional [15] untuk memastikan pertumbuhan dan keselamatan penerbangan sipil di seluruh dunia [1]. Hipotesis tersebut akan diuji kebenarannya dengan metode analisis yang digunakan.

### **3.2 Data Collection**

Pada tahap *data collection*, data kecelakaan pesawat terbang dunia yang dikumpulkan berasal dari *website* <https://www.baaa-acro.com/> [62] yang telah dikumpulkan oleh Ronan Hubert, seorang sejarawan dalam kecelakaan pesawat, dan ahli dalam kecelakaan penerbangan sejak tahun 1918. Data kecelakaan pesawat terbang dunia yang dikumpulkan adalah dari tahun 1918 – 2020 dengan *region world*. Data kecelakaan pesawat yang telah dikumpulkan akan dipindahkan kedalam sebuah file dengan ekstensi csv.

Data kecelakaan pesawat terbang BAAA tersebut awalnya berjumlah 15 kolom dan 592 baris, dimana tidak seluruh kolom dan baris data tersebut akan digunakan oleh penulis untuk keperluan analisis. Sehingga nantinya penulis akan mempersiapkan data mentah tersebut terlebih dahulu sebelum digunakan untuk keperluan analisis pada tahap *data preprocessing* selanjutnya [22].

### **3.3 Data Preprocessing**

Setelah mengumpulkan data kecelakaan pesawat terbang yang akan dianalisis, selanjutnya penulis akan melanjutkan ke tahap *data preprocessing* sebelum menganalisis data [22]. Pada penulisan ini, teknik preprocessing yang akan penulis gunakan adalah teknik *data transformation*, *data reduction* dan *data cleaning* untuk menyelesaikan masalah-masalah pada data mentah yang dimiliki untuk kebutuhan analisis selanjutnya [28].

### **3.3.1 Load Data**

Pertama, penulis akan memulai dengan melakukan *load* data mentah kecelakaan pesawat terbang BAAA yang berada pada file csv kedalam *software jupyter notebook* dengan bantuan library python yaitu *pandas*. Setelah *diload*, didapatkan bahwa data awal kecelakaan pesawat terbang BAAA memiliki 592 baris data dan 15 kolom dalam bentuk dataframe.

### **3.3.2 Renaming The Variable's Column Name**

Selanjutnya penulis akan mengganti judul dari setiap variabel agar lebih mudah untuk ditulis kedalam kode program. Disini penulis akan mengubah setiap karakter pada variabel menjadi huruf kecil (*lowercase*), dan menggantikan spasi dengan garis bawah (*underline*). Disini penulis tidak melakukan perubahan terhadap jumlah baris dan kolom data, sehingga jumlahnya tetap sama.

### **3.3.3 Rename The Variable's Name**

Selanjutnya, penulis akan membersihkan data dengan memeriksa apakah terdapat nilai yang hilang atau *null* pada data yang digunakan. Ketika terdapat nilai yang hilang, maka penulis akan menghapus baris data tersebut. Setelah memeriksa seluruh nilai pada data, tidak ditemukan nilai yang hilang, sehingga tidak ada baris data yang perlu dihapus atau diperbaiki, sehingga penulis dapat melanjutkan ke tahap selanjutnya.



### 3.3.4 Remove The Accidents Data of Non-Commercial

#### Flight

Kemudian, penulis akan menghapus baris data yang memiliki jenis penerbangan non-komersial. Perlu diketahui bahwa data kecelakaan pesawat terbang BAAA yang dikumpulkan sebelumnya, terdiri dari data kecelakaan pesawat terbang komersial dan non-komersial. Sementara pada penulisan ini, penulis hanya akan menganalisis data kecelakaan pesawat terbang komersial saja, sehingga baris data dengan jenis penerbangan non-komersial akan dihapus.

Berdasarkan penjelasan pada bab sebelumnya, ICAO mengklasifikasikan penerbangan *commercial* kedalam penerbangan *scheduled* dan *non-scheduled* [20]. Pada data kecelakaan pesawat terbang BAAA yang digunakan, variabel yang terdiri dari jenis penerbangan adalah variabel Flight Types. Pada variabel Flight Types, penerbangan *scheduled* dikategorikan dengan nama “scheduled revenue flight”, sementara penerbangan non scheduled dikategorikan dengan

nama “air/taxi”, “cargo”, dan “postal”. Setelah menghapus baris data dengan jenis penerbangan non-komersial, jumlah baris data yang awalnya berjumlah 592 baris, berkurang menjadi 110 baris saja.

### **3.3.5 Remove The Data which Contain Unknown Causes of Accidents**

Selanjutnya, penulis akan menghapus baris data yang memiliki penyebab kecelakaan yang tidak diketahui, dan dilabeli dengan “Unknown Reasons” pada variabel probable causes. Hal ini dikarenakan kecelakaan pesawat terbang komersial <sup>111</sup> dengan penyebab yang tidak jelas dan tidak diketahui akan sulit untuk dideskripsikan dan dianalisis. Setelah berhasil menghapus baris data yang memiliki penyebab kecelakaan yang tidak diketahui, terdapat pengurangan jumlah baris data dari 110 baris, menjadi 91 baris.

### 3.3.6 Variables Selection

Setelah melakukan beberapa tahapan *preprocessing*, terakhir penulis akan menentukan variabel-variabel mana saja yang akan digunakan untuk keperluan analisis data lebih lanjut. Variabel-variabel yang digunakan untuk keperluan analisis data adalah variabel tahun (*years*), variabel penyebab kecelakaan (*probable causes*), variabel jenis penerbangan (*flight types*) dan variabel total jumlah kematian (*total fatalities*). Sementara variabel lainnya yang tidak digunakan akan dihapus dari dataframe, sehingga dari 15 kolom, hanya menyisahkan 3 kolom variabel saja. Setelah melewati tahap *data preprocessing*, data kecelakaan pesawat terbang komersial BAAA yang dianalisis memiliki 91 baris dan 3 kolom.

### 3.4 Descriptive Statistics

Setelah melakukan persiapan terhadap data kecelakaan pesawat terbang komersial BAAA yang akan dianalisis pada tahap *data preprocessing*, selanjutnya penulis akan mendeskripsikan karakteristik dari setiap variabel pada data

tersebut dengan teknik *descriptive statistics*, sehingga data kecelakaan pesawat terbang komersial BAAA tersebut memiliki informasi yang berguna bagi pengguna data [32].

Sesuai penjelasan pada bab sebelumnya, terdapat beberapa jenis pengukuran untuk mendeskripsikan suatu data, dimana untuk dapat memilih jenis pengukuran yang tepat, penulis perlu mengetahui jenis data pada setiap variabel yang akan dideskripsikan. Setelah mengetahui jenis data dari masing-masing variabel, barulah penulis dapat menentukan jenis pengukuran yang tepat untuk mendeskripsikan karakteristik masing-masing variabel tersebut.

#### **3.4.1 Quantitative Data**

Pada data kecelakaan pesawat terbang komersial BAAA, terdapat dua buah variabel yang memiliki jenis data *quantitative*, yaitu variabel tahun (*years*) dan variabel total kematian (*total fatalities*). Sesuai penjelasan pada bab sebelumnya, terdapat beberapa pengukuran untuk mendeskripsikan variabel dengan

jenis data *quantitative*, yaitu *measures of frequency*, *measures of central tendency*, *measures of dispersion*, *skewness* dan *kurtosis* [35].

#### **3.4.1.1 Measures of Frequency**

Pada *measures of frequency*, penulis akan menghitung frekuensi kemunculan setiap nilai atau kategori yang berbeda tersebut pada data untuk melihat seberapa sering nilai atau kategori tersebut muncul pada data kecelakaan pesawat terbang yang digunakan [36]. Untuk dapat memvisualisasikan frekuensi kemunculan setiap nilai atau kategori tersebut, penulis dapat menggunakan tabulasi atau grafik diagram batang.

#### **3.4.1.2 Measures of Central Tendency**

Setelah mendeskripsikan kedua variabel tersebut menggunakan pengukuran frekuensi kemunculan pada masing-masing kategori nilainya, selanjutnya penulis akan mendeskripsikan kedua variabel tersebut dengan menentukan

lokasi titik pusat distribusi data pada setiap variabel. Terdapat tiga jenis parameter pengukuran untuk menentukan titik pusat dari distribusi data, yaitu dengan *mean*, *median* dan *mode* [38]. Sebelum dapat memvisualisasikan grafik distribusi data pada masing-masing variabel, penulis perlu menghitung seluruh parameter pengukuran titik pusat distribusi data terlebih dahulu.

Setelah memvisualisasikan distribusi data pada setiap variabel, jika variabel yang digunakan memiliki jenis distribusi data simetris, maka penulis dapat memilih salah satu nilai dari parameter *mean*, *median* atau *mode*. Sedangkan jika suatu variabel memiliki distribusi data *skewed*, maka penulis hanya dapat menggunakan parameter *median* untuk menentukan titik pusat distribusi data, dimana parameter median tidak memiliki pengaruh yang besar terhadap *outliers* [39].

#### **3.4.1.3 Measures of Dispersion**

Setelah menentukan titik pusat distribusi data pada kedua variabel yang digunakan, selanjutnya penulis akan menentukan

seberapa jauh menyebar setiap nilai dari titik pusat distribusi data masing-masing variabel tersebut. Terdapat beberapa parameter pengukuran untuk menentukan persebaran data dari titik pusat distribusi data, yaitu *range*, IQR, *variance*, dan *standard deviation* [44]. Sebelum dapat memilih jenis parameter pengukuran persebaran data yang tepat, penulis perlu menghitung seluruh parameter pengukuran persebaran data tersebut pada masing-masing variabel, lalu akan memvisualisasikan persebaran data berdasarkan parameter pengukuran yang digunakan. Sama seperti pada pengukuran titik pusat distribusi data sebelumnya, tidak seluruh parameter pengukuran persebaran data tersebut akan digunakan untuk menentukan persebaran nilai dari titik pusat distribusi data, tetapi bergantung dari jenis distribusi data yang dimiliki masing-masing variabel.

Ketika suatu variabel memiliki jenis distribusi data simetris, maka penulis dapat menggunakan parameter pengukuran *variance* atau *standard deviation* yang

dikombinasikan dengan nilai *mean*. Sedangkan jika suatu variabel memiliki jenis distribusi data *skewed*, maka parameter pengukuran yang dapat digunakan adalah IQR, dimana IQR tidak memiliki pengaruh yang besar terhadap outliers. Berdasarkan visualisasi persebaran data tersebut, ketika setiap nilai menyebar terlalu jauh dari titik pusat distribusi data, maka data kecelakaan pesawat terbang komersial BAAA tersebut memiliki lebih banyak nilai outliers dan setiap nilai tidak mirip satu sama lainnya. Sementara ketika setiap nilai mengelompok di sekitar titik pusat data, maka setiap nilai pada data lebih konsisten dan memiliki lebih sedikit outliers [44].

#### **3.4.1.4 Skewness dan Kurtosis**

Setelah mengukur persebaran data dari titik pusat distribusi datanya, penulis akan menghitung nilai dari parameter *skewness* dan *kurtosis* untuk mengukur kurva distribusi data setiap variabel.



Pertama, penulis akan menghitung nilai *skewness* untuk mengukur seberapa miring kurva distribusi data yang dimiliki pada setiap variabel, dari kurva distribusi data yang normal. Ketika nilai *skewness* nol, maka variabel tersebut memiliki jenis distribusi data simetris (*symmetrical distribution*). Sementara ketika nilai *skewness* positif atau berada di *range* positif, maka jenis distribusi data yang dimiliki adalah *positive skewed*, yang ditandai dengan kurva distribusi data yang miring ke arah kanan. Sedangkan ketika *skewness* bernilai negatif atau berada di *range* negatif, maka jenis distribusi data yang dimiliki adalah *negative skewed*, yang ditandai dengan kurva distribusi data yang miring ke arah kiri [52].

Selanjutnya penulis akan menghitung nilai kurtosis untuk menentukan jenis distribusi data berdasarkan puncak dan ketebalan ekor kurva distribusi data setiap variabel. Seperti yang telah dijelaskan sebelumnya terdapat 3 jenis distribusi data berdasarkan nilai kurtosis, yaitu *mesokurtic*, *leptokurtic*, dan *platykurtic* [54]. *Mesokurtic* merupakan jenis distribusi data

normal yang memiliki nilai kurtosis 3. *Platykurtic* merupakan jenis distribusi data yang memiliki nilai kurtosis negatif atau kurang dari 3, dengan ekor kurva yang tipis dan memiliki *outliers* dengan jumlah yang sedikit. Sedangkan *leptokurtic* merupakan jenis distribusi data yang memiliki nilai kurtosis lebih besar dari 3, dengan ekor kurva yang tebal, serta memiliki lebih banyak outliers [53].

### 3.4.2 Qualitative Data

Setelah mendeskripsikan variabel dengan jenis data *quantitative*, selanjutnya penulis akan mendeskripsikan variabel yang memiliki jenis data *qualitative*. Pada data kecelakaan pesawat terbang komersial BAAA yang digunakan, terdapat satu variabel yang memiliki jenis data *qualitative*, yaitu variabel penyebab kecelakaan (*probable causes*). Pada variabel dengan jenis data *qualitative* ini, penulis akan mendeskripsikan variabel tersebut dengan menghitung frekuensi kemunculan dan persentase atau proporsi dari masing-masing kategori penyebab

kecelakaan pada data kecelakaan pesawat terbang komersial BAAA yang digunakan [36].

### **3.5 Exploratory Data Analysis (EDA)**

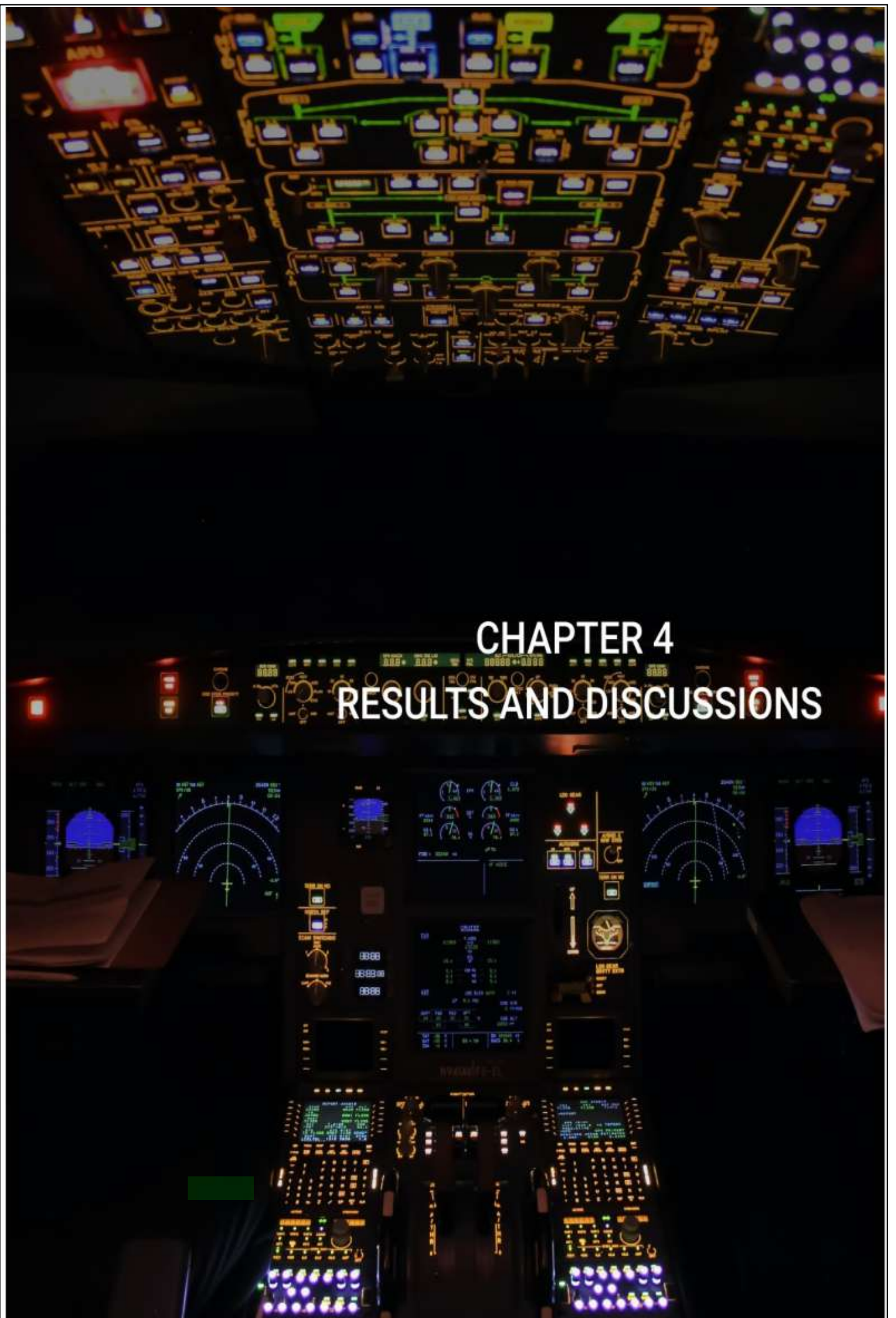
Setelah mendeskripsikan karakteristik setiap variabel pada data kecelakaan pesawat terbang komersial BAAA dengan menggunakan teknik *descriptive statistics*, penulis telah mengetahui karakteristik dari masing-masing variabel tersebut seperti frekuensi kemunculan setiap kategori data, titik pusat distribusi data, persebaran data dari titik pusat distribusi data, dan lainnya yang akan digunakan sebagai salah satu landasan dalam menganalisis data yang digunakan dengan metode EDA.

Disini penulis akan menggunakan metode EDA untuk menemukan relasi dari setiap variabel yang digunakan, yang berfokus kepada jumlah dan tingkat kecelakaan dan kematian pesawat terbang komersial selama periode tahun 1918 – 2020. Setiap relasi tersebut akan divisualisasikan dalam bentuk diagram untuk dilakukan analisis. Penulis akan menganalisis

perubahan pada grafik jumlah dan tingkat kecelakaan dan kematian pesawat terbang komersial dalam dua bagian, yaitu sebelum ICAO didirikan (periode tahun 1918 – 1947) dan setelah ICAO didirikan (periode tahun 1948 – 2020), sehingga penulis dapat melihat apakah terdapat perubahan terhadap grafik kecelakaan dan kematian pesawat terbang komersial setelah ICAO dibentuk.

44

*This Page Intentionally Left Blank*



# CHAPTER 4 RESULTS AND DISCUSSIONS

*This Page Intentionally Left Blank*



## CHAPTER 4

### RESULTS AND DISCUSSIONS

#### 4.1 Descriptive Statistics

Setelah melewati tahap data preprocessing, penulis akan melakukan analisis data dengan melakukan deskripsi terhadap setiap variabel yang digunakan pada data. Metode analisis pertama yang penulis akan digunakan adalah *descriptive statistics*. Sesuai penjelasan pada bab sebelumnya, terdapat beberapa pengukuran yang bisa digunakan untuk mendeskripsikan sebuah data, tetapi bergantung kepada jenis data tersebut.

##### 4.1.1 The Variables with Quantitative Data Type

Pada data kecelakaan pesawat terbang komersial yang akan dianalisis, terdapat dua variabel dengan jenis data



quantitative yaitu variabel jumlah kematian akibat kecelakaan pesawat terbang komersial (*total fatalities*), dan tahun terjadinya kecelakaan pesawat terbang komersial (*years*). Jenis pengukuran yang penulis gunakan untuk mendeskripsikan kedua variabel tersebut adalah *measures of frequency*, *measures of central tendency*, *measures of dispersion*, serta *skewness* dan *kurtosis* [35].

#### 4.1.1.1 Measures of Frequency

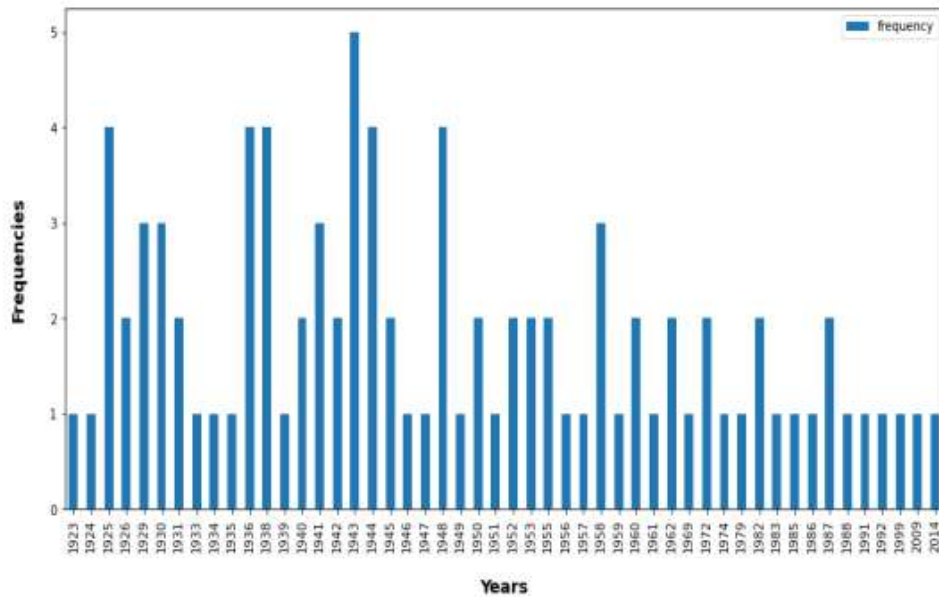
Pertama, penulis akan mendeskripsikan kedua variabel menggunakan pengukuran frekuensi kemunculan setiap nilai yang berbeda pada data. Pengukuran frekuensi kemunculan ini digunakan untuk melihat berapa kali sebuah nilai atau kategori tertentu muncul pada data kecelakaan pesawat terbang komersial BAAA yang digunakan.

Variabel pertama yang akan dideskripsikan adalah *years*. Variabel ini berisi data tahun terjadinya kecelakaan pesawat komersial dari tahun 1918-2020. Penulis akan

mengukur berapa kali kategori tahun berbeda muncul pada data.

Hasil pengukuran frekuensi kemunculan pada variabel years

ditunjukkan <sup>58</sup> dalam bentuk diagram batang pada **Gambar 4.1**.



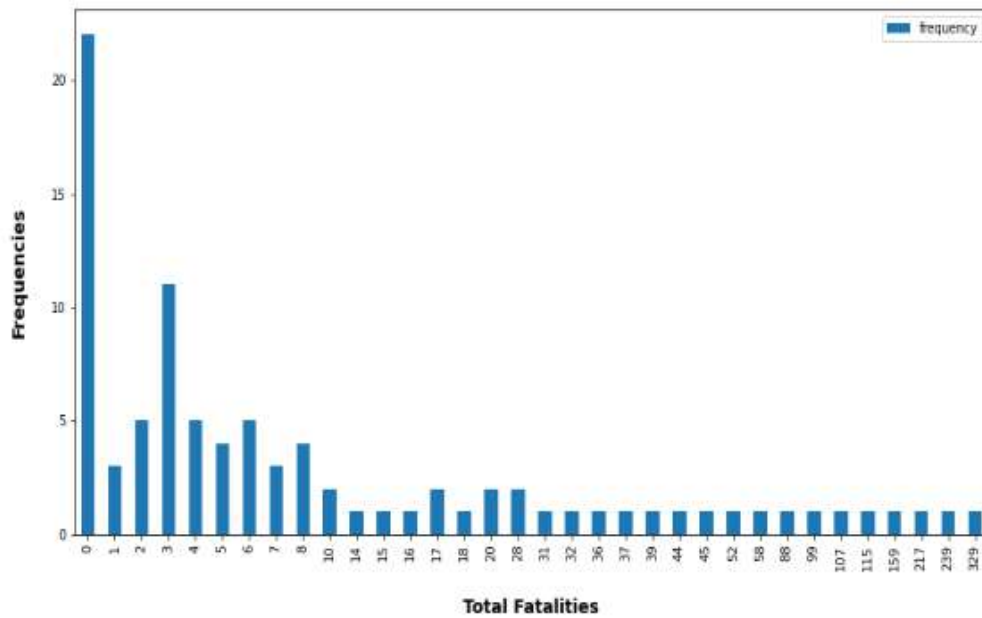
**Gambar 4.1** Visualisasi pengukuran frekuensi kemunculan pada variabel *years*

Dari diagram batang tersebut, didapatkan bahwa tahun 1943 merupakan kategori tahun dengan jumlah kemunculan paling banyak pada data, sebanyak 5 kali. Kemudian dilanjutkan dengan tahun 1925, 1936, 1938, 1944, dan 1948 dengan jumlah kemunculan terbanyak kedua, sebanyak 4 kali kemunculan.

Selanjutnya, kategori tahun dengan jumlah kemunculan tertinggi ketiga sebanyak 3 kali kemunculan adalah tahun 1929, 1930, 1941, dan 1958. Kategori tahun dengan jumlah kemunculan tertinggi keempat sebanyak 2 kali adalah tahun 1950, 1960, 1952, 1953, 1955, 1940, 1942, 1962, 1945, 1972, 1982, 1931, 1987, dan 1926. Sementara kategori tahun sisanya seperti tahun 1923, 1924, 1933, 1934, 1935, 1939, 1946, 1947, 1949, 1951, 1956, 1957, 1959, 1961, 1969, 1974, 1979, 1983, 1985, 1986, 1988, 1991, 1992, 1999, 2009, 2014 memiliki jumlah kemunculan paling sedikit pada data yang digunakan, yaitu sebanyak 1 kali kemunculan.

Variabel selanjutnya yang akan dideskripsikan dengan menghitung jumlah frekuensi kemunculan setiap nilai nya adalah *total fatalities*. Variabel ini terdiri dari jumlah kematian dalam satu kasus kecelakaan pesawat terbang komersial. Disini penulis akan menghitung jumlah kemunculan setiap angka kematian berbeda <sup>100</sup> pada data, yang ditunjukkan pada **Gambar**

**4.2.**



**Gambar 4.2** Visualisasi pengukuran frekuensi kemunculan pada variabel *total fatalities*

Dari diagram frekuensi kemunculan setiap angka kematian akibat kecelakaan pesawat terbang komersial, dapat dilihat bahwa kasus kecelakaan pesawat terbang komersial dengan jumlah kematian sebanyak 0 jiwa memiliki jumlah kemunculan paling banyak, dengan 22 kali kemunculan. Kemudian kasus kecelakaan pesawat terbang komersial dengan jumlah kematian sebanyak 3 jiwa memiliki jumlah kemunculan paling banyak kedua dengan 11 kali kemunculan. Selanjutnya

kecelakaan dengan jumlah kematian sebanyak 2, 4, dan 6 jiwa memiliki jumlah kemunculan paling banyak ketiga sebanyak 5 kali pada data. Dilanjutkan dengan jumlah kematian sebanyak 5 dan 8 jiwa memiliki jumlah kemunculan pada data sebanyak 4 kali. Lalu kecelakaan dengan kematian sebanyak 1 dan 7 jiwa memiliki kemunculan 3 kali, serta untuk kecelakaan dengan angka kematian 10, 17, 20, dan 28 jiwa muncul sebanyak 2 kali pada data. Sementara untuk angka kematian lainnya seperti 14, 15, 16, 18, 31, 32, 36, 37, 39, 44, 45, 52, 58, 88, 99, 107, 115, 159, 217, 239, 329 muncul sebanyak 1 kali pada data.

#### **4.1.1.2 Measures of Central Tendency**

Setelah mendeskripsikan kedua variabel dengan mengukur frekuensi kemunculan setiap nilainya, selanjutnya penulis akan mendeskripsikan variabel tersebut dengan mengukur letak titik pusat distribusi data kedua variabel tersebut dengan pengukuran pemusatan data atau *measures of central tendency*. Pada pengukuran pemusatan data, terdapat beberapa

parameter pengukuran yang digunakan yaitu *mean*, *median*, dan *mode* [38]. Sesuai penjelasan pada bab sebelumnya, untuk dapat memilih parameter pengukuran yang tepat untuk menentukan titik pusat distribusi data, penulis perlu mengetahui jenis distribusi data pada setiap variabel yang akan digunakan.

Pertama, penulis akan mengukur pemusatan data pada variabel *years*. Sesuai penjelasan sebelumnya, penulis akan menghitung nilai *mean*, *median* dan *mode* pada variabel *years*. Hasil perhitungan ketiga parameter pengukuran pemusatan data tersebut ditunjukkan dalam bentuk tabel <sup>24</sup> pada **Tabel 4.1**. Dari tabel tersebut, dapat dilihat bahwa nilai *mean* dari variabel *years* adalah 1951.098, nilai *median* adalah 1945, dan nilai *mode* adalah 1943.

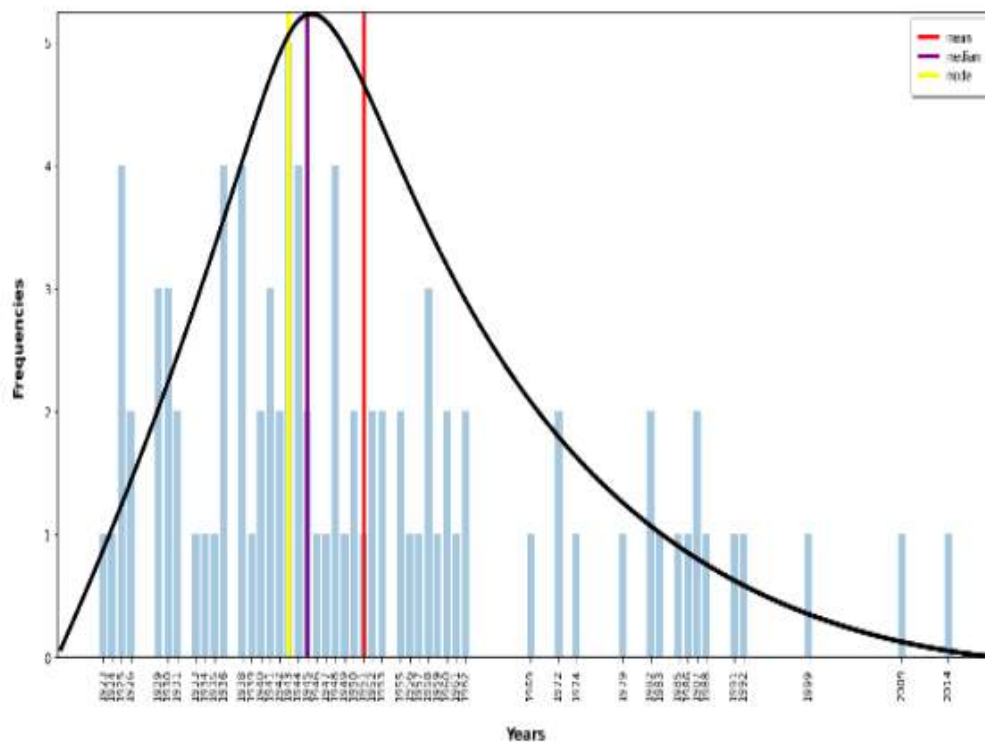
**Tabel 4.1** Hasil perhitungan parameter pengukuran pemusatan data pada variabel *years*

No	Measures of Central Tendency Parameters	Values
1	Mean	1951.098
2	Median	1945
3	Mode	1943

Setelah menghitung seluruh parameter pengukuran pemusatan data, untuk dapat memilih parameter pengukuran pemusatan yang tepat untuk menentukan titik pusat dari variabel year, penulis harus mengetahui jenis distribusi data terlebih dahulu dengan menggambarkan kurva distribusi data berdasarkan nilai *mean*, *median*, dan *mode* seperti pada

102

**Gambar 4.3.**



**Gambar 4.3** Visualisasi pemusatan data pada variabel *years*



Dari kurva distribusi data tersebut, dapat dilihat bahwa nilai mean berada di sebelah kanan dari nilai median, dan nilai median berada di sebelah kanan nilai mode. Dalam hal ini dapat diartikan bahwa variabel year memiliki jenis distribusi data tidak simetris (*skewed*), dan tergolong kedalam jenis distribusi data miring positif (*positive skewed*) berdasarkan letak ketiga parameter tersebut pada diagram [63]. Ketika suatu variabel memiliki jenis distribusi data *skewed*, maka parameter pengukuran yang dapat dipilih untuk menentukan titik pusat distribusi data adalah *median*. Hal ini dikarenakan distribusi data *skewed* memiliki nilai ekstrim atau *outliers* didalamnya, dan dari ketiga parameter pemusatan data tersebut, hanya parameter *median* tidak memiliki pengaruh yang besar terhadap *outliers* [39]. Oleh karena itu, titik pusat distribusi data pada variabel *years* adalah 1945.

Selanjutnya penulis akan menentukan letak titik pusat distribusi data pada variabel *total fatalities*. Sama seperti pada



variabel *years* sebelumnya, penulis akan menghitung nilai dari parameter *mean*, *median*, dan *mode* terlebih dahulu. Hasil perhitungan ketiga parameter tersebut ditunjukkan pada Tabel 4.2. Pada tabel tersebut, dapat dilihat bahwa nilai *mean* dari variabel *total fatalities* adalah 23.175, nilai *median* adalah 4.0, dan nilai *mode* adalah 0.

**Tabel 4.2** Perhitungan parameter pemusatan data pada variabel *total fatalities*

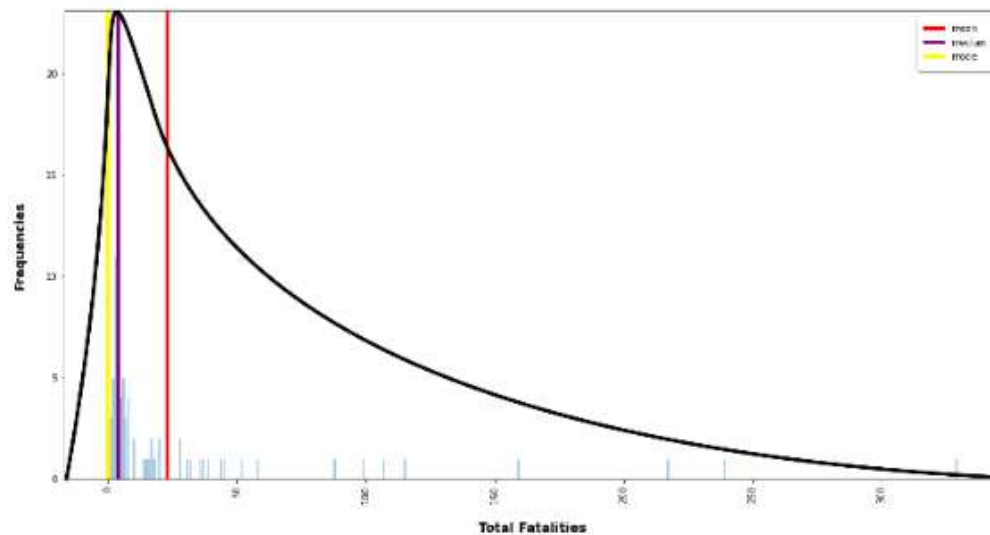
No	Measures of Central Tendency Parameters	Values
1	Mean	23.175
2	Median	4.0
3	Mode	0

Setelah menghitung seluruh parameter pemusatan data, selanjutnya penulis akan menampilkan distribusi data pada variabel *total fatalities* berdasarkan nilai *mean*, *median*, dan *mode*. Dengan mengetahui jenis distribusi data, maka penulis dapat memilih parameter pengukuran mana yang tepat untuk menentukan titik pusat distribusi data pada variabel *total*

*fatalities*. Visualisasi distribusi data pada variabel total fatalities

86

dapat dilihat pada **Gambar 4.4**.



**Gambar 4.4** Visualisasi pemusatan data pada variabel *total fatalities*

Terlihat bahwa letak nilai *mean*, *median*, dan *mode* pada grafik distribusi data variabel total fatalities mirip dengan variabel *years* sebelumnya, dimana nilai *mean* berada di sebelah kanan dari nilai *median*, dan nilai *median* berada di sebelah kanan nilai *mode*. Hal ini menunjukkan bahwa variabel total fatalities juga memiliki jenis distribusi data *positive skewed*.

Karena memiliki jenis distribusi data skewed, maka penulis akan menggunakan nilai dari parameter *median* untuk menentukan titik pusat distribusi data, karena median tidak memiliki pengaruh yang besar terhadap *outliers* [39]. Sehingga nilai titik pusat distribusi data pada variabel *total fatalities* adalah 4.0.

#### **4.1.1.3 Measures of Dispersion**

Setelah menentukan titik pusat distribusi data pada kedua variabel yang digunakan, selanjutnya penulis akan menentukan seberapa jauh persebaran setiap titik data dari titik pusat distribusi data kedua variabel tersebut.

Variabel pertama yang akan diukur persebaran data nya adalah variabel *years*. Sebelumnya, penulis telah menentukan titik pusat distribusi data pada variabel *years* dengan menggunakan nilai parameter median, yaitu 1945. Selanjutnya penulis akan mengukur seberapa menyebar setiap titik data pada variabel *years* terhadap titik pusat distribusi datanya dengan menggunakan parameter pengukuran penyebaran setiap titik

data seperti *range*, IQR, *variance* dan *standard deviation* [42].

Hasil perhitungan setiap parameter penyebaran data pada variabel *years* terlihat pada Tabel 4.3.

**Tabel 4.3** Perhitungan parameter dispersi data pada variabel *years*

No	Measures of Dispersion Parameters	Values
1	Range	91
2	Interquartile Range (IQR)	22.5
3	Population Variance	423.606
4	Standard Deviation	20.582

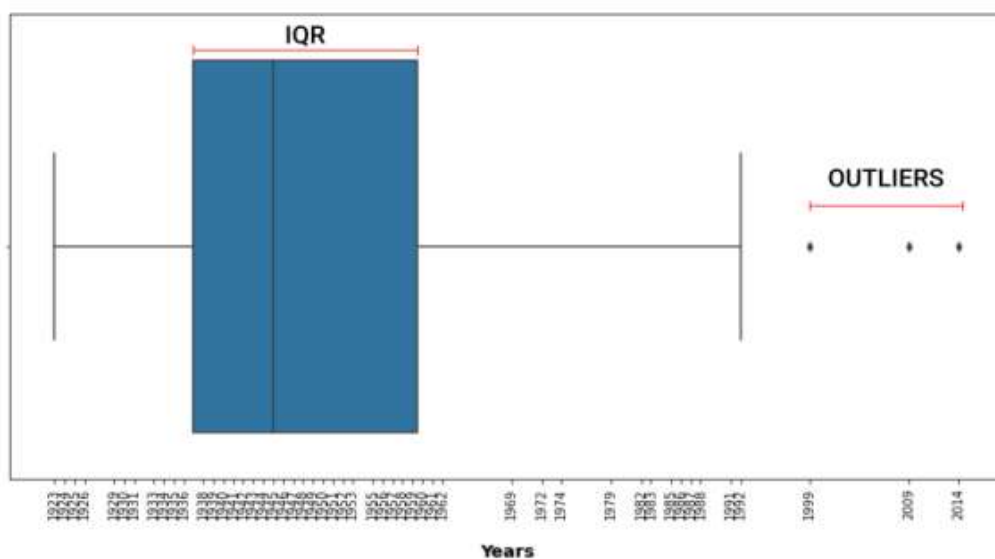
Setelah menghitung seluruh parameter pengukuran penyebaran data yang tersedia, selanjutnya penulis akan memilih parameter mana yang tepat untuk menentukan penyebaran data pada variabel *years*. Pada perhitungan pemusatan data sebelumnya, diketahui bahwa variabel *year* memiliki jenis distribusi *skewed*, yang memiliki beberapa nilai outliers [39]. Sesuai penjelasan pada bab sebelumnya, parameter pengukuran penyebaran data seperti *range*, *variance*, dan *standard deviation* sangat sensitif dengan adanya *outliers*.

Sementara parameter IQR memiliki sifat yang sama dengan *median*, yang tidak memiliki pengaruh besar terhadap nilai *outliers* [44]. Sehingga penulis akan menggunakan IQR sebagai parameter pengukuran penyebaran data variabel *years*. Grafik persebaran data dengan menggunakan IQR dapat dilihat pada

**Gambar 4.5.**

Dari grafik persebaran data pada variabel *years* menggunakan parameter IQR, dapat dilihat bahwa kotak berwarna biru tersebut merupakan jarak persebaran titik data normal, yang dimulai dari Q1 dengan nilai 1937, dan diakhiri dengan Q3 dengan nilai 1959.5. Sementara titik pusat distribusi datanya ditandai dengan garis hitam yang berada di tengah kotak berwarna biru tersebut dengan nilai 1945 yang didapatkan dari pengukuran titik pusat distribusi data sebelumnya. Disini IQR merupakan selisih antara Q3 dan Q1, sehingga parameter IQR memiliki nilai 22.5. Selain itu, terdapat batas minimum dan batas maksimum persebaran data yang ditoleransi, dimana batas minimumnya adalah 1923, dan batas maksimumnya adalah

1992. Titik data yang berada diluar batas minimum dan maksimum yang ditentukan akan dianggap sebagai *outliers* atau nilai ekstrim. Pada diagram terlihat bahwa variabel *years* memiliki 3 nilai yang dianggap sebagai outliers, yaitu 1999, 2009, dan 2014.



**Gambar 4.5** Visualisasi persebaran titik data dari titik pusat distribusi data pada variabel *years*

Setelah mengukur persebaran data dari titik pusat pada variabel *years*, selanjutnya penulis akan mengukur persebaran setiap titik data dari titik pusat distribusi datanya pada variabel

*total fatalities*. Sama seperti pada variabel *years*, penulis akan menghitung seluruh nilai parameter pengukuran penyebaran data pada variabel *total fatalities*. Hasil perhitungan parameter penyebaran data tersebut dapat dilihat dalam bentuk tabel pada **Tabel 4.4**.

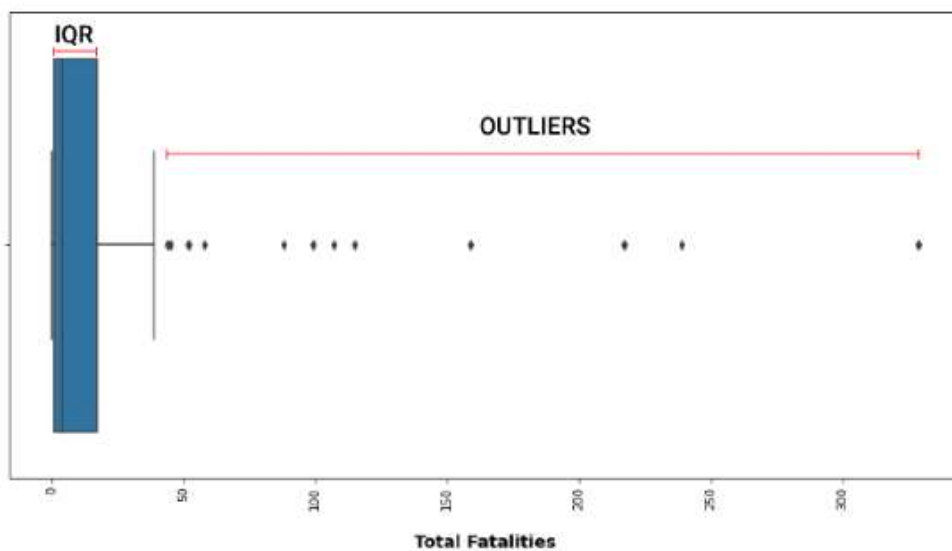
**Tabel 4.4** Perhitungan parameter dispersi data pada variabel *total fatalities*

No	Measures of Dispersion Parameters	Values
1	Range	329
2	Interquartile Range (IQR)	16.0
3	Population Variance	2772.584
4	Population Standard Deviation	52.655

Setelah menghitung seluruh parameter pengukuran penyebaran data pada variabel *total fatalities*, sama seperti pada variabel *years* sebelumnya, penulis perlu mengetahui jenis distribusi data variabel tersebut terlebih dahulu sebelum menentukan parameter pengukuran mana yang akan digunakan. Sesuai hasil pengukuran pemusatan data sebelumnya, penulis mengetahui bahwa variabel *total fatalities* memiliki jenis



distribusi data *skewed*. Dikarenakan jenis distribusi data adalah *skewed*, maka parameter pengukuran penyebaran data yang akan digunakan untuk menentukan penyebaran setiap titik data dari titik pusat distribusi data variabel *total fatalities* adalah IQR. Hal ini dikarenakan parameter IQR tidak memiliki pengaruh besar terhadap nilai outliers [44]. Visualisasi penyebaran data pada variabel *total fatalities* menggunakan parameter IQR terlihat pada Gambar 4.6.



**Gambar 4.6** Visualisasi persebaran titik data dari titik pusat distribusi data pada variabel *total fatalities*



Berdasarkan diagram yang menunjukkan penyebaran data pada variabel total fatalities, dapat dilihat bahwa persebaran data normal pada variabel total fatalities ditunjukkan dengan sebuah kotak berwarna biru, yang dimulai dari Q1 yang bernilai 1, hingga Q3 yang memiliki nilai 17. Dikarenakan IQR merupakan selisih antara Q3 dan Q1, maka didapatkan nilai IQR sebesar 16. Selain itu dapat dilihat bahwa terdapat batas minimum dan maksimum persebaran data pada variabel total fatalities, dimana <sup>128</sup> batas minimumnya adalah 0, dan batas maksimumnya adalah 41. Setiap nilai yang berada diluar batas minimum dan maksimum tersebut dianggap sebagai *outliers*. Pada variabel *total fatalities*, terdapat lebih banyak *outliers* dibandingkan variabel *years* sebelumnya. Seluruh angka kematian dengan *range* antara 42 hingga 329 dianggap sebagai *outliers*.

#### 4.1.1.4 Skewness and Kurtosis

Setelah menghitung frekuensi kemunculan setiap nilai atau kategori berbeda, menentukan titik pusat distribusi data, dan menentukan seberapa jauh persebaran setiap titik data dari titik pusat distribusi data setiap variabel, selanjutnya penulis akan mendeskripsikan kedua variabel tersebut dengan mengukur kurva distribusi data menggunakan parameter pengukuran *skewness* dan *kurtosis*.

Variabel pertama yang akan diukur adalah variabel *years*. Disini penulis akan mengukur kurva distribusi data pada variabel *years* dengan menghitung nilai dari parameter *skewness* dan *kurtosis*. Hasil perhitungan *skewness* dan *kurtosis* pada variabel year <sup>24</sup> ditunjukkan pada **Tabel 4.5**.

**Tabel 4.5** Perhitungan parameter *skewness* dan *kurtosis* pada variabel *years*

No	Skewness and kurtosis parameters	Values
1	Skewness	1.015
2	Kurtosis	0.518

Setelah menghitung nilai *skewness* dan *kurtosis*, selanjutnya penulis akan mendeskripsikan variabel *years* dengan masing-masing parameter tersebut. Pertama, penulis akan mendeskripsikan variabel *years* dengan parameter *skewness*. Pada **Tabel 4.3**, nilai *skewness* yang didapatkan adalah 1.015. Sesuai penjelasan pada bab sebelumnya, ketika nilai *skewness* berada kurang dari -1 atau lebih dari 1, maka jenis distribusi data adalah sangat miring (*highly skewed*) [51]. Setelah mengetahui jenis distribusi data, selanjutnya penulis akan melihat arah kemiringan dari kurva distribusi data tersebut. Karena nilai *skewness* adalah positif, maka jenis kemiringan kurva distribusi data adalah *positive skewed* atau miring ke arah kanan [52].

Kemudian untuk parameter *kurtosis*, didapatkan hasilnya adalah 0.518. Sesuai penjelasan pada bab sebelumnya, nilai *kurtosis* untuk distribusi normal adalah 3. Sehingga *kurtosis* yang didapat pada variabel *years* kurang dari distribusi normal, atau dalam kata lain nilai *kurtosis* yang didapatkan bernilai

negatif. Kurtosis dengan nilai negatif memiliki jenis kurva *platykurtic*, yang ditandai dengan ekor kurva yang tipis (*light-tailed*), dan terdapat sedikit *outliers* [53]. Hal ini sesuai dengan persebaran data variabel *years* pada **Gambar 4.5**, dimana hanya sedikit *outliers* yang dimiliki.

Setelah mendeskripsikan kurva distribusi data pada variabel *years*, selanjutnya penulis akan melakukan hal yang sama pada variabel total fatalities. Pada variabel *total fatalities*, penulis juga akan menghitung nilai *skewness* dan *kurtosis*, yang hasilnya dapat dilihat pada **Tabel 4.6**.

**Tabel 4.6** Perhitungan parameter *skewness* dan *kurtosis* pada variabel *total fatalities*

No	Skewness and kurtosis parameters	Values
1	Skewness	3.830
2	Kurtosis	16.324

Setelah menghitung nilai *skewness* dan *kurtosis*, selanjutnya penulis akan mendeskripsikan variabel *total fatalities* dengan masing-masing parameter tersebut. Pertama

adalah *skewness*, dari hasil perhitungan didapatkan nilai *skewness* adalah 3.830. Sesuai penjelasan pada bab sebelumnya, ketika nilai *skewness* berada kurang dari -1 atau lebih dari 1, maka jenis distribusi data pada variabel *total fatalities* adalah sangat miring (*highly skewed*) [51]. Lalu dikarenakan nilai *skewness* yang didapatkan bernilai positif, maka jenis kemiringan kurva distribusi data adalah *positive skewed* atau miring ke arah kanan [52]. Kemudian berdasarkan perhitungan sebelumnya, didapatkan nilai *kurtosis* sebesar 16.324. Ketika nilai *kurtosis* yang didapatkan lebih besar dari nilai *kurtosis* pada distribusi data normal, maka dapat diartikan bahwa nilai *kurtosis* yang didapatkan memiliki nilai positif. Ketika *kurtosis* bernilai positif, maka variabel tersebut memiliki jenis kurva distribusi data *leptokurtic*, dimana kurva ini memiliki ekor yang tebal (*heavy-tailed*), dan memiliki lebih banyak nilai yang berada di ekor kurva yang mengindikasikan bahwa distribusi data memiliki lebih banyak nilai *outliers* [53]. Hal ini sesuai dengan grafik penyebaran data pada **Gambar 4.6**,

dimana terdapat banyak nilai yang berada di ekor kurva dibandingkan yang berada disekitar titik pusat distribusi data.

#### **4.1.2 The Variables with Qualitative Data Type**

Setelah mendeskripsikan variabel dengan jenis data *quantitative*, selanjutnya penulis akan mendeskripsikan variabel dengan jenis data *qualitative*. Pada data kecelakaan pesawat terbang komersial yang dianalisis, terdapat dua variabel dengan jenis data *qualitative*, yaitu variabel penyebab kecelakaan (*probable causes*) dan jenis penerbangan (*flight types*). Jenis pengukuran yang penulis gunakan untuk mendeskripsikan kedua variabel tersebut adalah menghitung frekuensi kemunculan, dan persentase dari masing-masing kategori nya pada data [36].

##### **4.1.2.1 Flight Types Variable**

Variabel pertama yang akan dideskripsikan adalah variabel *flight types*. Variabel ini terdiri dari jenis-jenis aktivitas penerbangan komersial berdasarkan pembagian oleh ICAO

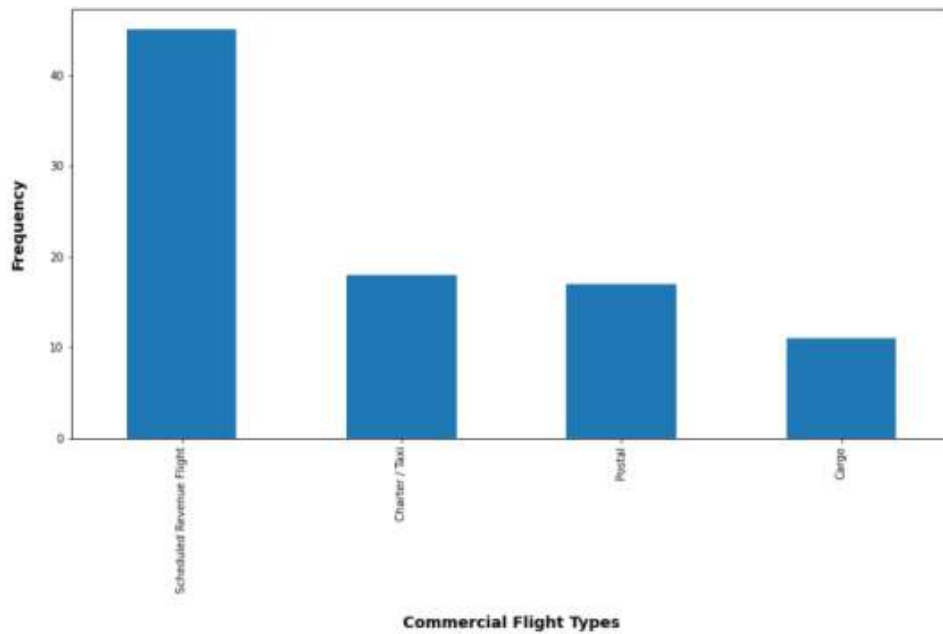
seperti yang terlihat pada **Gambar 2.1**. Pada data yang digunakan, terdapat empat jenis aktivitas penerbangan komersial, yaitu *scheduled revenue flight*, *cargo*, *postal*, dan *charter / taxi flight*. Disini penulis akan menghitung frekuensi kemunculan dan persentase dari setiap jenis aktivitas penerbangan komersial tersebut pada data yang digunakan.

Pertama, penulis akan menghitung frekuensi kemunculan setiap kategori penyebab kecelakaan pada data yang digunakan. Hasil perhitungannya divisualisasikan dalam bentuk diagram batang pada **Gambar 4.7**.

Berdasarkan diagram batang diatas, dapat dilihat bahwa kategori jenis aktivitas penerbangan komersial *Scheduled Revenue Flight* memiliki frekuensi kemunculan tertinggi sebanyak 45 kali kemunculan. Kemudian kategori aktivitas penerbangan komersial dengan jumlah frekuensi tertinggi kedua adalah *Charter / Taxi* sebanyak 18 kali kemunculan. Lalu, kategori jenis aktivitas penerbangan komersial *Postal* sebanyak 17 kali kemunculan, serta kategori jenis aktivtias penerbangan



komersial dengan frekuensi kemunculan terendah adalah *Cargo* sebanyak 11 kali kemunculan.



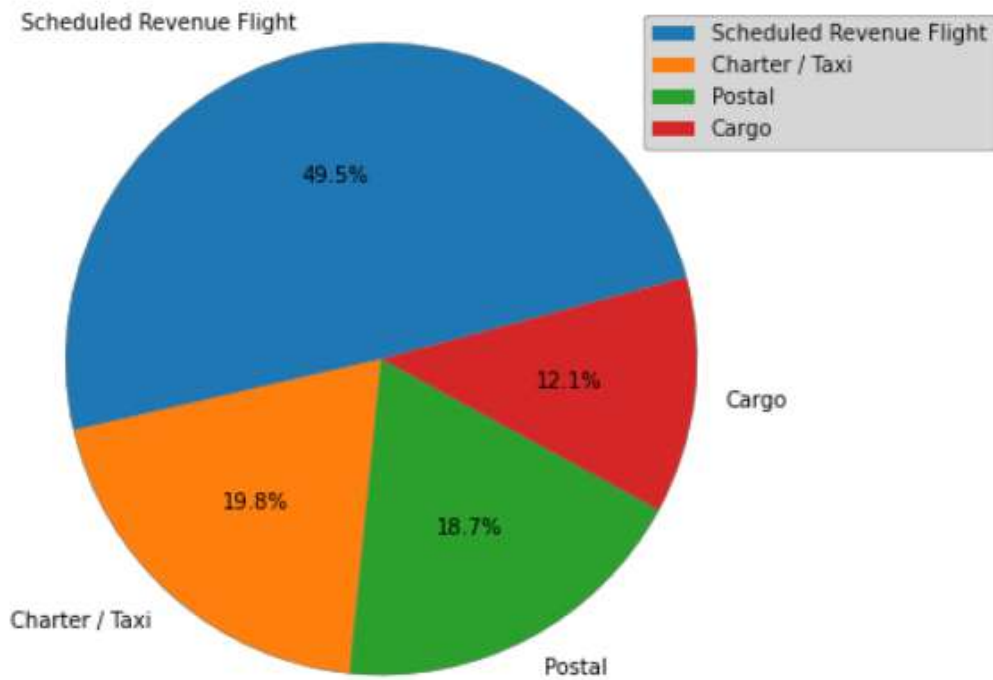
**Gambar 4.7** Visualisasi frekuensi kemunculan masing-masing aktivitas penerbangan komersial

Setelah menghitung frekuensi kemunculan setiap kategori aktivitas penerbangan komersial, selanjutnya penulis akan menghitung persentase dari masing-masing kategori tersebut,



yang <sup>4</sup> divisualisasikan dalam bentuk diagram lingkaran pada

**Gambar 4.8.**



**Gambar 4.8** Visualisasi persentase masing-masing aktivitas penerbangan komersial pada data

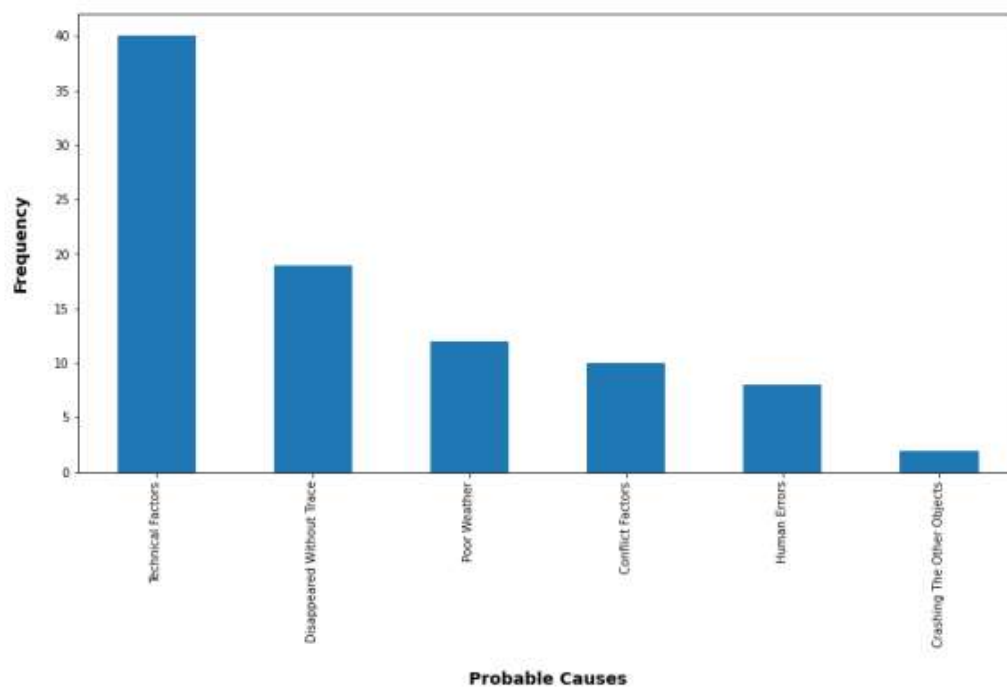
Berdasarkan diagram lingkaran tersebut, dapat dilihat bahwa kategori aktivitas penerbangan komersial Scheduled Revenue Flight memiliki persentase tertinggi pada data sebesar 49.5%. Kemudian dilanjutkan dengan kategori aktivitas

penerbangan komersial Charter / Taxi dengan persentase tertinggi kedua sebesar 19.8%, lebih rendah 29.7% dibandingkan kategori Scheuled Revenue Flight. Lalu kategori akitvitas penerbangan komersial Postal dengan persentase tertinggi ketiga sebesar 18.7%, dan kategori aktivitas penerbangan komersial Cargo dengan persentase terendah pada data, sebesar 12.1%.

#### **4.1.2.2 Probable Causes Variable**

Variabel dengan jenis data *qualitative* selanjutnya yang akan penulis deskripsikan adalah variabel *probable causes*. Variabel ini terdiri dari penyebab terjadinya kecelakaan pesawat terbang komersial pada data yang digunakan. Sesuai penjelasan pada bab sebelumnya, penulis akan mendeskripsikan variabel dengan menghitung frekuensi kemunculan, dan persentase dari masing-masing kategori penyebab kecelakaan pada data yang digunakan.

Pertama, penulis akan menghitung frekuensi kemuculan setiap kategori penyebab kecelakaan pada variabel *probable causes*. Hasil perhitungan frekuensi kemunculan setiap kategori penyebab kecelakaan dapat dilihat pada **Gambar 4.9** dalam bentuk diagram batang.



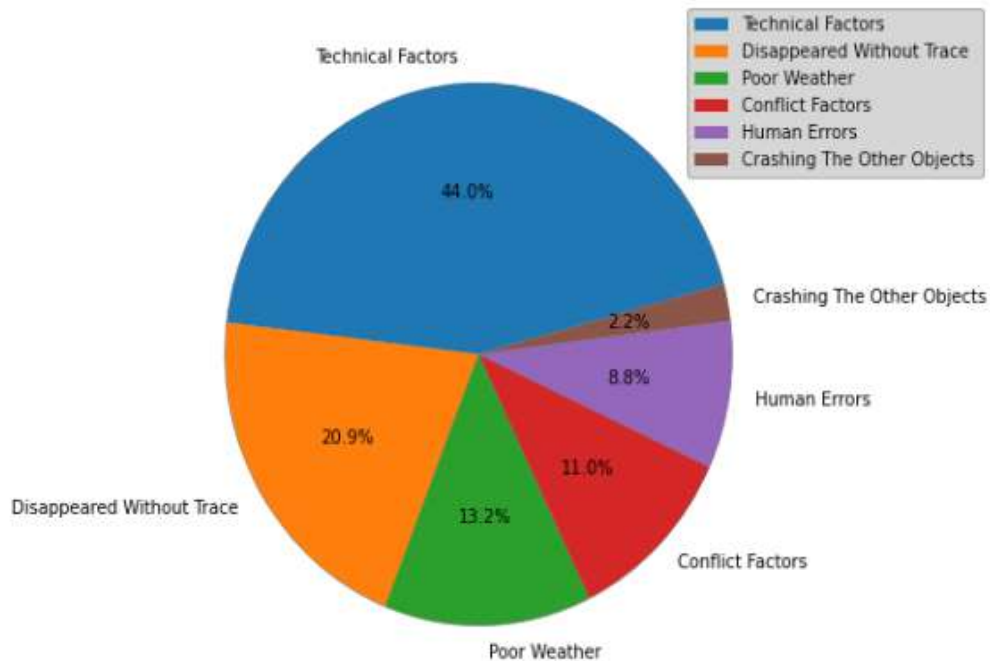
**Gambar 4.9** Frequency Distribution pada variabel Probable Cause

Terlihat bahwa kategori penyebab kecelakaan “Technical Factors” memiliki frekuensi kemunculan tertinggi pada data sebanyak 40 kali kemunculan. Kemudian kategori penyebab kecelakaan “Dissapeared Without Trace” memiliki frekuensi kemunculan tertinggi kedua sebanyak 19 kali kemunculan. Selanjutnya, kategori penyebab kecelakaan dengan frekuensi kemunculan tertinggi ketiga adalah “Poor Weather” dengan kemunculan sebanyak 12 kali. Kategori penyebab kecelakaan tertinggi keempat adalah “Conflict Factors” dengan kemunculan sebanyak 10 kali, lalu yang kelima adalah kategori “Human Errors” dengan 8 kali kemunculan. Sementara kategori penyebab kecelakaan dengan frekuensi kemunculan paling rendah adalah “Crashing the Other Objects” dengan kemunculan sebanyak 2 kali pada data.

Setelah mendeskripsikan variabel *probable causes* dengan menggunakan perhitungan frekuensi kemunculan setiap kategori, selanjutnya penulis akan menghitung persentase dari setiap kategori penyebab kecelakaan terhadap keseluruhan

kategori pada variabel *probable causes*. Hasil perhitungan persentase dapat dilihat dalam bentuk diagram lingkaran pada

**Gambar 4.10.**



**Gambar 4.10** Persentase kategori penyebab kecelakaan pada variabel *probable causes*

Dari diagram lingkaran gambar 4.10, diketahui bahwa 44% dari jumlah seluruh kecelakaan pesawat terbang komersial pada data yang digunakan disebabkan oleh faktor teknis

“Technical Factors”. Kemudian kategori penyebab kecelakaan “Dissapeared Without Trace” memiliki persentase tertinggi kedua sebesar 20.9%. Selanjutnya kategori penyebab kecelakaan “Poor Weather” memiliki persentase tertinggi ketiga sebesar 13.2%. Faktor konflik “Conflict Factors” dengan persentase sebesar 11.0%, 8.8% disebabkan oleh faktor kesalahan manusia “Human Errors”, dan kategori penyebab kecelakaan “Crashing The Other Objects” dengan persentase terendah sebesar 2.2%.

#### **4.2 Exploratory Data Analysis**

Setelah mendeskripsikan karakteristik dari masing-masing variabel yang digunakan dengan metode *descriptive statistics*, penulis akan menggunakan hasil deskripsi tersebut sebagai salah satu referensi dalam menganalisis data kecelakaan pesawat terbang komersial yang digunakan dengan metode EDA. Disini penulis akan menggunakan metode EDA untuk menemukan relasi antara setiap variabel yang digunakan, yang berfokus kepada jumlah kecelakaan dan kematian pesawat

terbang komersial per tahun, selama periode tahun 1918–2020. Masing-masing relasi tersebut akan divisualisasikan dalam bentuk diagram, baik diagram garis, batang, atau lingkaran untuk dianalisis lebih lanjut. Dalam menganalisis perubahan grafik kecelakaan dan kematian pesawat terbang komersial pada masing-masing tersebut, penulis akan membagi analisis kedalam dua bagian, yaitu jumlah kecelakaan dan kematian pesawat terbang komersial per tahun pada periode tahun 1918-1947 sebelum ICAO dibentuk, serta pada periode tahun 1948-2020 setelah ICAO dibentuk.

#### **4.2.1 Number of Accidents and Fatalities per Year**

Pertama, penulis akan menganalisis jumlah kecelakaan pesawat terbang komersial per tahun selama periode tahun 1918-2020. Disini penulis akan memvisualisasikan grafiknya dengan diagram garis seperti pada **Gambar 4.11**. Berdasarkan diagram garis tersebut, dapat dilihat bahwa jumlah kecelakaan pesawat terbang komersial per tahun paling banyak terjadi pada



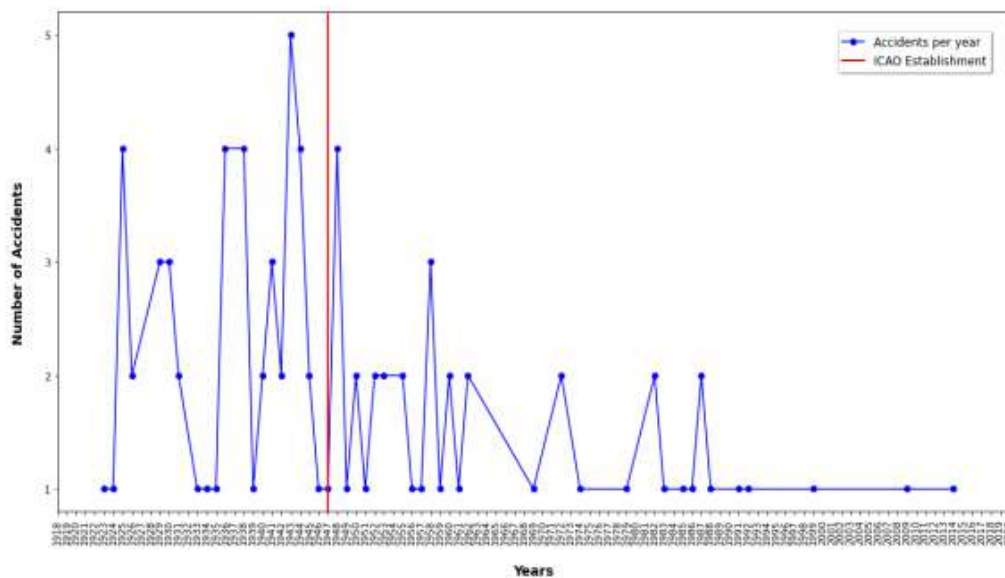
tahun 1943 sebanyak 5 kasus kecelakaan. Sementara jumlah kecelakaan pesawat terbang komersial per tahun paling sedikit terjadi sebanyak 1 kali. Kemudian dapat dilihat bahwa pada periode tahun 1918-1947 sebelum dibentuk ICAO, grafik dan intensitas jumlah kecelakaan pesawat terbang komersial per tahun masih sangat tinggi, dimana hampir setiap tahun selalu mengalami kecelakaan, dan gap antar tahun juga kecil. Sementara setelah dibentuknya ICAO, grafik dan intensitas jumlah kecelakaan pesawat terbang komersial per tahun mengalami penurunan yang cukup drastis, dimana dapat dilihat bahwa semakin bertambahnya tahun, semakin terbentuk gap yang cukup besar antara setiap tahun yang mengalami kecelakaan, seperti antara tahun 1961 dan 1969, atau 1999 dan 2009. Berdasarkan diagram tersebut juga dapat dilihat bahwa selama periode tahun 1918-1947 sebelum dibentuknya ICAO, dengan jarak 29 tahun, telah terjadi kecelakaan pesawat terbang komersial sebanyak 48 kali pada 21 kategori tahun berbeda. Sementara pada periode tahun 1948-2020 setelah dibentuknya



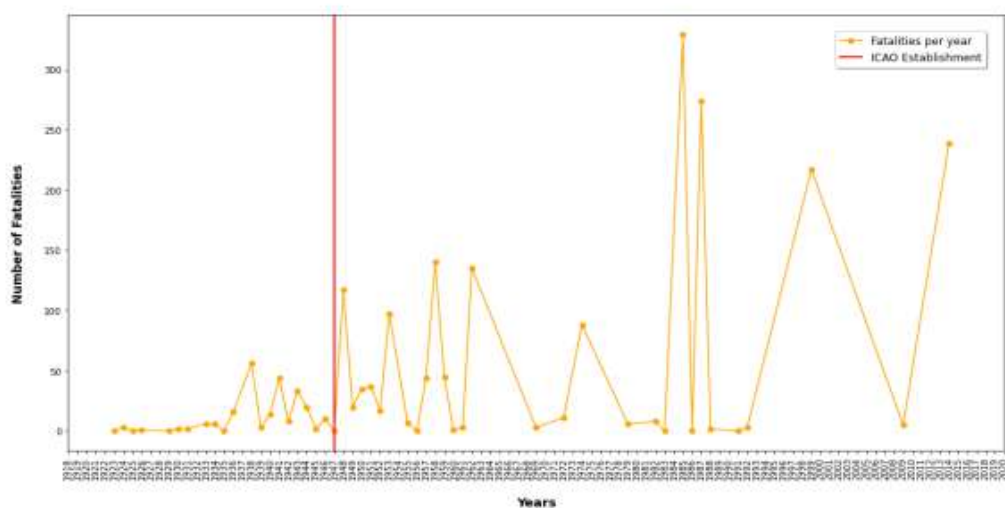
ICAO, dengan jarak 72 tahun, hanya mengalami kecelakaan pesawat terbang komersial sebanyak 42 kali pada 29 kategori tahun berbeda. Sehingga dapat dibuktikan bahwa setelah dibentuknya ICAO, terdapat penurunan grafik dan intensitas kecelakaan pesawat terbang komersial per tahun.

Selanjutnya, penulis akan menganalisis grafik jumlah kematian per tahun akibat kecelakaan pesawat terbang komersial selama periode tahun 1918-2020. Sama seperti pada grafik sebelumnya, penulis akan memvisualisasikan grafik ini dengan menggunakan diagram garis seperti pada **Gambar 4.12**. Berdasarkan diagram garis tersebut, dapat dilihat bahwa pada periode tahun 1948-2020 setelah dibentuknya ICAO, terjadi peningkatan grafik jumlah kematian per tahun yang sangat signifikan dibandingkan periode tahun 1918-1947 sebelum dibentuknya ICAO. Hal ini cukup menarik karena berdasarkan hasil analisis grafik jumlah kecelakaan pesawat terbang komersial per tahun sebelumnya, grafik dan intensitas kecelakaan pesawat terbang komersial per tahun justru

mengalami penurunan yang signifikan setelah dibentuknya ICAO. Pada periode tahun 1918-1947 sebelum dibentuknya ICAO, dengan jarak 29 tahun, telah terjadi 48 kasus kecelakaan dengan total jumlah kematian sebanyak 226 korban jiwa, dengan rata-rata kematian sebanyak 4-5 korban jiwa per satu kasus kecelakaan. Sementara pada periode tahun 1948-2020 setelah dibentuknya ICAO, dengan jarak 72 tahun, telah terjadi 43 kasus kecelakaan dengan total jumlah kematian sebanyak 1883 korban jiwa, dengan rata-rata kematian sebanyak 43-44 korban jiwa per satu kasus kecelakaan, jauh lebih tinggi dibandingkan sebelum dibentuknya ICAO, walaupun dengan jumlah kecelakaan yang lebih sedikit.



**Gambar 4.11** Visualisasi grafik jumlah kecelakaan pesawat terbang komersial per tahun selama periode tahun 1918-2020.



**Gambar 4.12** Visualisasi grafik jumlah kematian pesawat terbang komersial per tahun selama periode tahun 1918-2020

#### **4.2.2 Number of Accidents and Fatalities per Year by**

##### **Flight Types**

Selanjutnya, penulis akan menganalisis grafik jumlah kecelakaan dan kematian per tahun berdasarkan jenis aktivitas penerbangan komersial. Walaupun data yang digunakan adalah hanya data kecelakaan pesawat terbang komersial, namun tidak seluruh jenis aktivitas penerbangan komersial bertujuan membawa penumpang, sehingga rasanya kurang relevan untuk membandingkan jumlah kematian pada pesawat yang memiliki tujuan membawa penumpang, dan pesawat yang hanya bertujuan mengantar barang atau surat. Sehingga penulis melakukan analisis grafik jumlah kecelakaan dan kematian per tahun dengan lebih spesifik lagi, yaitu berdasarkan jenis aktivitas penerbangannya. Diketahui bahwa terdapat 4 jenis aktivitas penerbangan komersial pada data, dengan masing-masing persentasenya dapat dilihat pada **Gambar 4.8**. Sesuai hasil analisis pada bagian sebelumnya, diketahui bahwa secara general, grafik dan intensitas kecelakaan pesawat terbang

komersial per tahun mengalami penurunan setelah dibentuknya ICAO. Sementara grafik kematian per tahun akibat kecelakaan pesawat terbang komersial meningkat walaupun jumlah kecelakaan berkurang. Disini penulis akan menganalisis grafik kecelakaan dan kematian per tahun pada masing-masing jenis aktivitas penerbangan komersial untuk melihat titik puncak dari grafik jumlah kecelakaan dan kematian pesawat terbang komersial per tahun, dan menentukan jenis aktivitas penerbangan komersial dengan jumlah kecelakaan dan kematian tertinggi pada tahun tertentu. Selain itu penulis juga akan membandingkan perubahan grafik dari setiap jenis aktivitas penerbangan komersial yang ada.

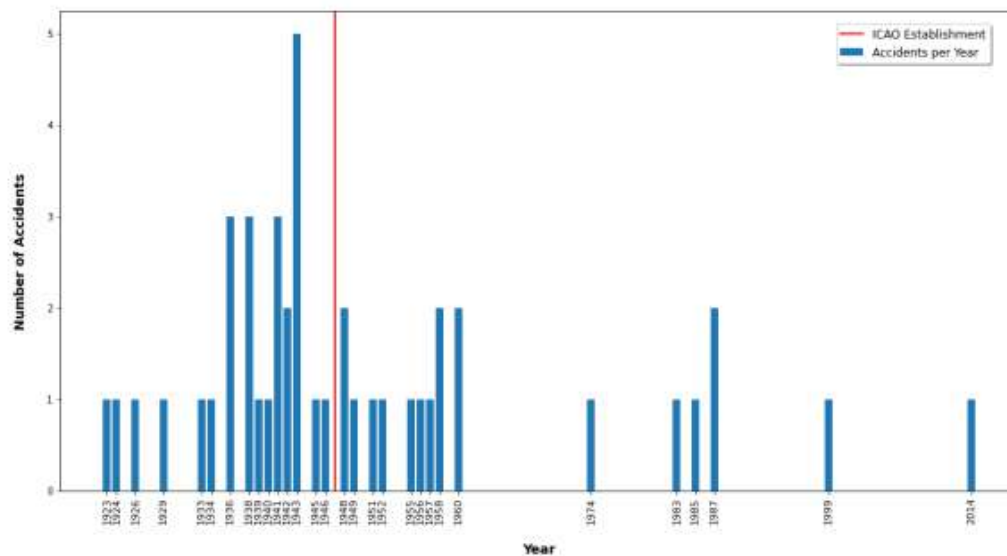
#### **4.2.2.1 Scheduled Revenue Flight**

Jenis aktivitas penerbangan komersial pertama adalah *scheduled revenue flight*. Disini penulis akan menganalisis grafik jumlah kecelakaan dan kematian pesawat terbang komersial dengan jenis aktivitas penerbangan *scheduled revenue*

*flight* yang divisualisasikan dalam bentuk diagram batang.

Visualisasi grafik jumlah kecelakaan pesawat terbang komersial

per tahun ditunjukkan pada **Gambar 4.13**.



**Gambar 4.13** Visualisasi jumlah kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan *scheduled revenue flight* per tahun

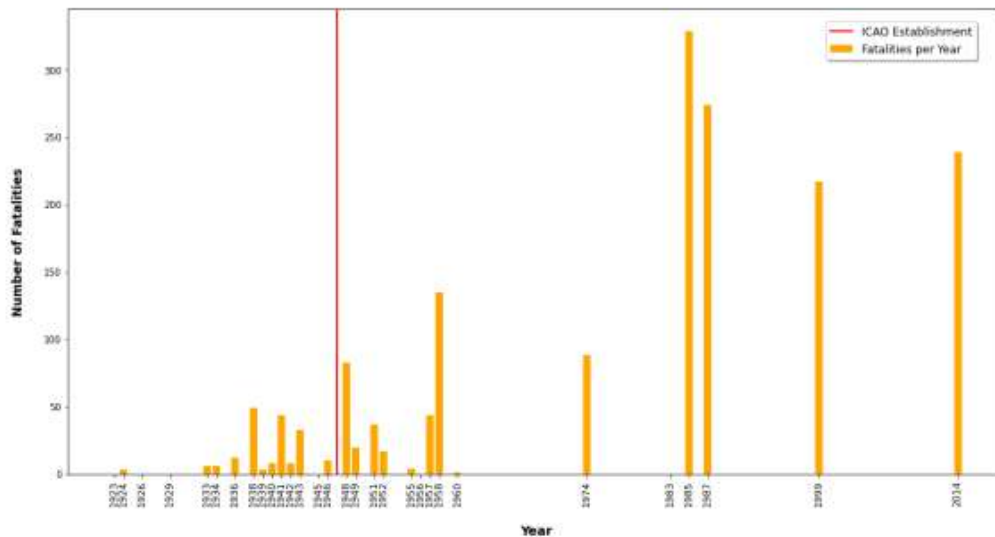
Berdasarkan diagram batang tersebut, dapat dilihat bahwa telah terjadi kecelakaan pesawat terbang komersial sebanyak 45 kali, dimana 26 kasus kecelakaan terjadi pada periode tahun 1918-1947 sebelum dibentuk ICAO, dan 19 kasus

sisanya terjadi pada periode tahun 1948-2020 setelah dibentuknya ICAO. Lalu, dapat dilihat pada grafik bahwa titik puncak pada grafik berada pada tahun 1943, dengan jumlah kecelakaan sebanyak 5 kali. Kemudian setelah dibentuk ICAO, grafik perlahan mengalami penurunan jumlah dan intensitas kecelakaan per tahun, yang dibuktikan dengan terjadinya kelonggaran antara jarak satu tahun dengan tahun lainnya. Sehingga dapat dikatakan bahwa dibentuknya ICAO sukses menurunkan grafik jumlah kecelakaan pesawat terbang komersial scheduled revenue flight dan intensitasnya per tahun. Selanjutnya penulis akan menganalisis grafik jumlah kematian per tahun yang ditunjukkan pada **Gambar 4.14**.

Berdasarkan diagram batang yang menampilkan grafik jumlah kematian pesawat terbang komersial per tahun diatas, dapat dilihat bahwa grafik jumlah kematian per tahun justru berbanding terbalik dengan grafik jumlah kecelakaan sebelumnya, dimana grafik jumlah kematian mengalami peningkatan yang sangat tajam setelah dibentuknya ICAO,



walaupun jumlah kecelakaan mengalami penurunan. Titik puncak pada grafik tersebut berada pada tahun 1985 dengan jumlah kematian sebanyak 329 jiwa per tahun. Dari data yang digunakan, didapatkan bahwa selama periode tahun 1918-1947, telah terjadi 26 kasus kecelakaan dengan total kematian sebanyak 182 jiwa..



**Gambar 4.14** Visualisasi jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan *scheduled revenue flight* per tahun



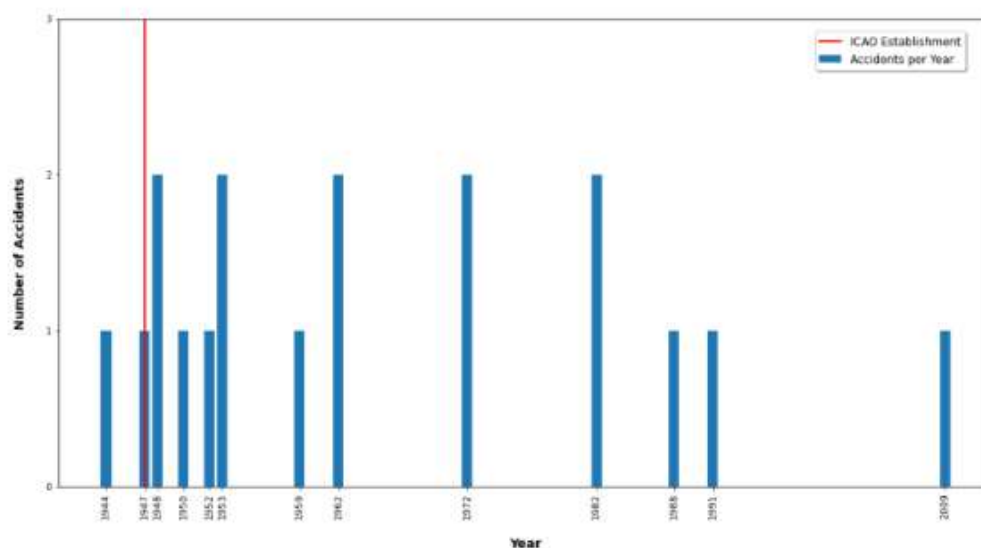
Dengan rata-rata kematian sebanyak 7 jiwa per satu kasus kecelakaan. Sedangkan selama periode tahun 1948-2020, telah terjadi 19 kasus kecelakaan dengan total kematian sebanyak 1488 jiwa. Dengan rata-rata kematian sebanyak 78 jiwa per satu kasus kecelakaan.

#### **4.2.2.2 Charter / Taxi Flight**

Jenis aktivitas penerbangan komersial selanjutnya adalah charter / taxi flight. Berdasarkan penjelasan pada bab sebelumnya, jenis aktivitas penerbangan ini termasuk kedalam non-scheduled commercial flight. Disini penulis akan menganalisis grafik kecelakaan dan kematian pesawat terbang komersial dengan aktivitas penerbangan charter / taxi flight per tahun selama periode tahun 1918-2020. Visualisasi grafik kecelakaan per tahun ditunjukkan pada **Gambar 4.15**.

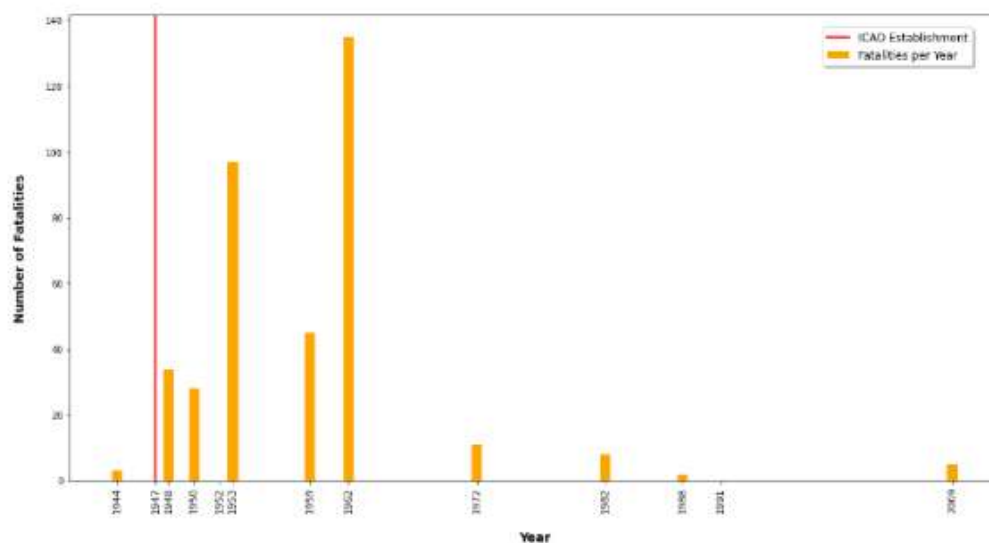
Berdasarkan diagram batang tersebut, dapat dilihat bahwa selama tahun 1918-2020 telah terjadi 18 kasus kecelakaan pada 12 kategori tahun berbeda, dimana 2 kasus

kecelakaan terjadi pada periode tahun 1918-1947 sebelum dibentuk ICAO, sedangkan 16 kasus kecelakaan sisanya terjadi pada periode tahun 1948-2020 setelah dibentuk ICAO. Selain itu, grafik jumlah kecelakaan per tahun juga mengalami peningkatan setelah dibentuknya ICAO. Hal ini menunjukkan bahwa pembentukan ICAO belum bisa mengurangi jumlah dan intensitas kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan *charter / taxi flight* per tahun.



**Gambar 4.15** Visualisasi jumlah kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan *charter / taxi flight* per tahun

Selanjutnya, penulis akan menganalisis jumlah kematian per tahun pesawat terbang komersial dengan jenis aktivitas penerbangan *charter / taxi flight* per tahun selama periode tahun 1918-2020 yang <sup>4</sup> divisualisasikan dalam bentuk diagram batang pada **Gambar 4.16**.



**Gambar 4.16** Visualisasi jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan charter / taxi flight per tahun

Berdasarkan diagram tersebut, dapat dilihat bahwa selama tahun 1918-2020, total kematian yang diakibatkan oleh

18 kasus kecelakaan pesawat terbang komersial charter / taxi adalah sebanyak 368 jiwa, dimana pada periode tahun 1918-1947 memiliki total kematian sebanyak 3 jiwa, dan pada periode tahun 1948-2020 memiliki total kematian sebanyak 365 jiwa. Sama seperti pada grafik jumlah kecelakaan per tahun sebelumnya, grafik jumlah kematian per tahun juga mengalami peningkatan setelah dibentuknya ICAO, dimana titik puncaknya adalah pada tahun 1962 dengan jumlah kematian sebanyak 132 jiwa dalam satu tahun. Peningkatan grafik jumlah kematian dirasa cukup masuk akal, karena grafik jumlah kecelakaan juga mengalami peningkatan. Walaupun begitu, grafik jumlah kematian per tahun mengalami penurunan secara drastis pada tahun 1972 hingga tahun 2009. Dengan ini, didapatkan bahwa pembentukan ICAO belum bisa mengurangi jumlah kematian akibat kecelakaan pesawat terbang komersial charter / taxi secara signifikan, hal ini dibuktikan dengan meningkatnya grafik jumlah kematian per tahun dibandingkan periode sebelumnya. Tetapi ICAO secara perlahan mulai memperbaiki

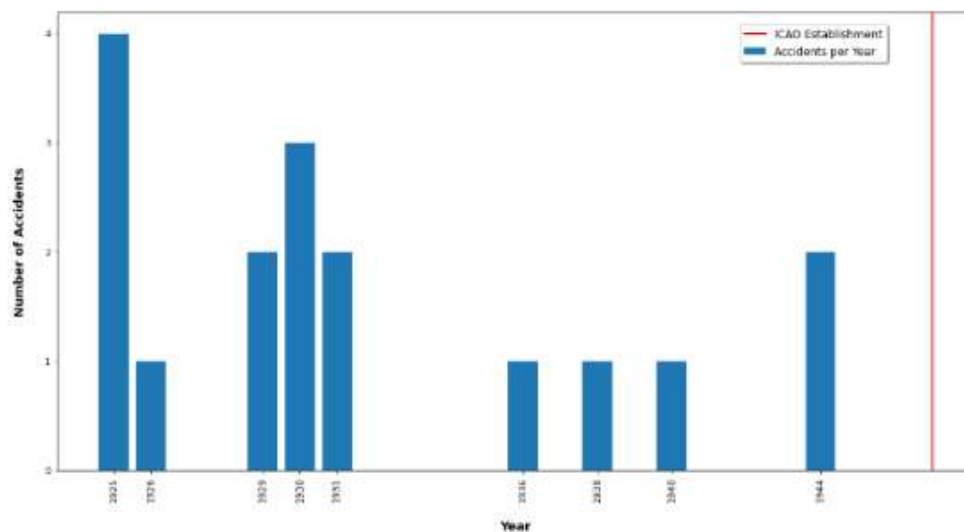
sistem keselamatan penerbangan mereka sehingga terjadi penurunan terhadap grafik jumlah kematian per tahun yang dimulai pada tahun 1972.

#### 4.2.2.3 Postal Flight

Jenis aktivitas penerbangan komersial selanjutnya adalah postal flight. Penerbangan ini bertujuan <sup>98</sup> untuk mengirimkan surat dari suatu tempat ke tempat lain. Seperti pada scheduled revenue flight sebelumnya, penulis akan menganalisis grafik jumlah kecelakaan dan kematian pesawat terbang komersial dengan jenis aktivitas penerbangan postal flight per tahun selama periode tahun 1918-2020. Visualisasi grafik jumlah kecelakaan pesawat terbang komersial per tahun ditunjukkan pada **Gambar 4.17**.

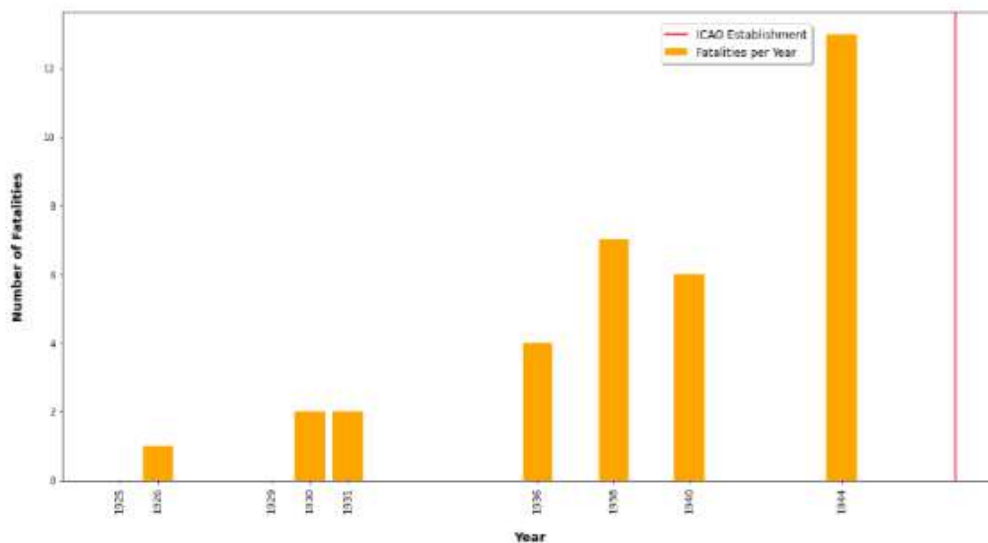
Berdasarkan diagram batang tersebut, dapat dilihat bahwa selama tahun 1918-2020, terjadi kecelakaan pesawat terbang komersial sebanyak 17 kali, dimana seluruh kecelakaan tersebut terjadi pada periode tahun 1918-1947 sebelum

dibentuknya ICAO. Titik puncak pada grafik berada pada tahun 1925 dengan kecelakaan sebanyak 4 kali dalam satu tahun. Dengan ini didapatkan bahwa, pembentukan ICAO berhasil mengurangi jumlah dan intensitas kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan postal flight per tahun.



**Gambar 4.17** Visualisasi jumlah kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan *postal flight* per tahun

Selanjutnya penulis akan menganalisis grafik jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan postal flight per tahun, yang <sup>4</sup> divisualisasikan dalam bentuk diagram batang pada **Gambar 4.18**.



**Gambar 4.18** Visualisasi jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan *postal flight* per tahun

Berdasarkan diagram batang tersebut, didapatkan bahwa selama periode tahun 1918-2020, total kematian akibat kecelakaan pesawat terbang postal komersial flight adalah

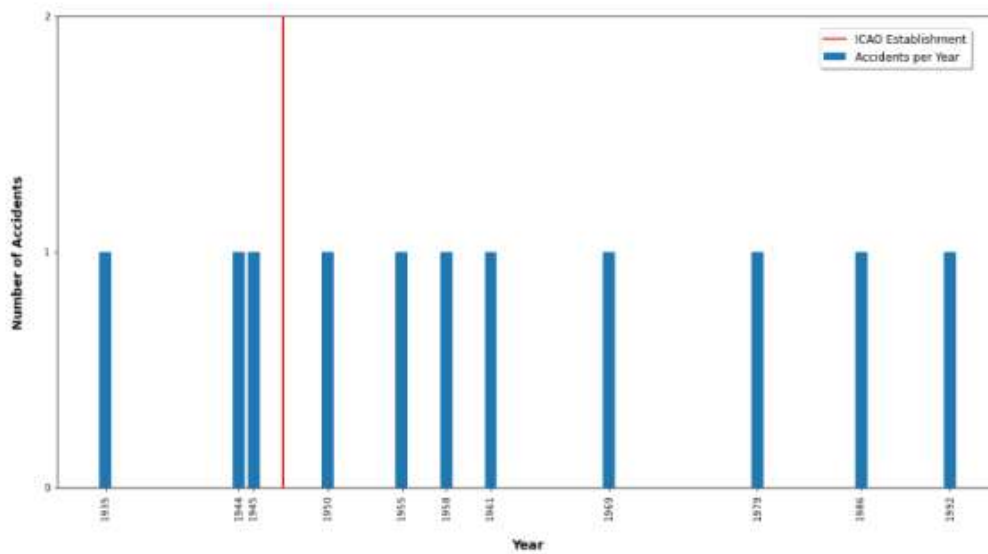


sebanyak 35 jiwa. Berdasarkan hasil analisis grafik jumlah kecelakaan pesawat terbang komersial postal flight sebelumnya, didapatkan bahwa total kematian tersebut dihasilkan dari 17 kasus kecelakaan selama periode tahun 1918-1947, sebelum dibentuknya ICAO. Dikarenakan setelah dibentuknya ICAO tidak terjadi satu pun kasus kecelakaan pesawat terbang komersial postal flight, maka jumlah kematian selama periode tahun 1948-2020 adalah sebanyak nol jiwa.

#### **4.2.2.4 Cargo Flight**

Jenis aktivitas penerbangan komersial selanjutnya adalah cargo flight. Cargo flight adalah aktivitas penerbangan komersial yang mengangkut barang. Penulis akan menganalisis grafik jumlah kecelakaan dan kematian pesawat terbang komersial dengan jenis aktivitas penerbangan komersial cargo flight per tahun selama periode tahun 1918-2020. Visualisasi grafik jumlah kecelakaan pesawat terbang komersial cargo <sup>2</sup> ditunjukkan pada **Gambar 4.19**.





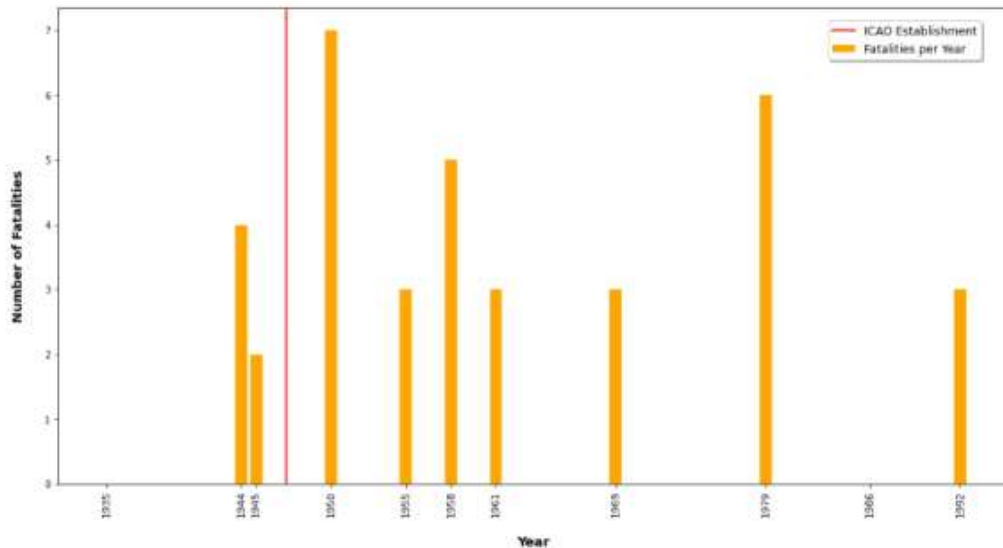
**Gambar 4.19** Visualisasi jumlah kecelakaan pesawat terbang komersial dengan jenis aktivitas penerbangan *cargo flight* per tahun

Berdasarkan diagram batang tersebut, didapatkan bahwa selama periode tahun 1918-2020, telah terjadi kecelakaan pesawat terbang komersial cargo sebanyak 11 kali pada 11 kategori tahun yang berbeda, dimana sebanyak 3 kasus kecelakaan terjadi pada periode tahun 1918-1947, sementara 8 kasus kecelakaan sisanya terjadi pada periode tahun 1948-2020. Dapat dilihat pada grafik tersebut bahwa tidak ada peningkatan

dan penurunan grafik. Hal ini dikarenakan seluruh kecelakaan terjadi pada kategori tahun yang berbeda, sehingga tidak ada satu kategori tahun yang memiliki jumlah kecelakaan lebih banyak dari kategori tahun lainnya. Tetapi, jika dihitung berdasarkan rata-rata kecelakaan, didapatkan bahwa pada periode tahun 1918-1947 sebelum dibentuk ICAO, dengan jarak 29 tahun, rata-rata kecelakaan sebanyak 9-10 kasus kecelakaan per tahunnya. Sementara pada periode tahun 1948-2020 setelah dibentuk ICAO, dengan jarak 72 tahun, didapatkan rata-rata kecelakaan sebanyak 9 kasus per tahunnya. Dengan ini didapatkan bahwa baik sebelum dan setelah dibentuknya ICAO, tidak ada penurunan dan kenaikan rata-rata kasus kecelakaan per tahunnya. Tetapi intensitas kecelakaan per tahun jelas mengalami penurunan setelah dibentuknya ICAO, dimana dapat dilihat pada grafik bahwa terdapat gap yang besar antar tahun.

Selanjutnya penulis akan menganalisis grafik jumlah kematian pesawat terbang komersial cargo per tahun selama

periode tahun 1918-2020, yang visualisasikan dalam bentuk diagram batang seperti pada Gambar 4.20.



**Gambar 4.20** Visualisasi jumlah kematian pesawat terbang komersial dengan jenis aktivitas penerbangan cargo flight per tahun

65

Berdasarkan diagram batang tersebut, dapat dilihat bahwa terjadi peningkatan grafik jumlah kematian per tahun setelah dibentuk ICAO. Hal ini dibuktikan dengan pada periode tahun 1918-1947 sebelum dibentuk ICAO, total kematian akibat 3 kasus kecelakaan pesawat terbang komersial cargo adalah

sebanyak 6 jiwa, dimana puncaknya pada tahun 1944 sebanyak 4 kasus kecelakaan per tahun, dan rata-rata kematian sebanyak 2 jiwa per 1 kasus kecelakaan. Sedangkan pada periode tahun 1948-2020 setelah dibentuk ICAO, total kematian yang diakibatkan oleh 8 kasus kecelakaan pesawat terbang komersial cargo adalah sebanyak 30 jiwa, dimana puncaknya pada tahun 1950 sebanyak 7 kematian dalam satu tahun, dengan rata-rata kematian sebanyak 3 hingga 4 jiwa per 1 kasus kecelakaan. Dengan ini didapatkan bahwa terjadi peningkatan jumlah dan rata-rata kematian setelah dibentuknya ICAO, walaupun intensitas kecelakaan per tahun mengalami penurunan sekalipun.

#### **4.2.3 Number of Accidents and Fatalities per Year by**

##### **Probable Causes**

Setelah melakukan analisis terhadap grafik jumlah kecelakaan dan kematian pesawat terbang komersial per tahun berdasarkan jenis aktivitas penerbangan, selanjutnya penulis akan menganalisis grafik tersebut berdasarkan masing-masing

kategori penyebab kecelakaan. Pada tahap analisis ini, penulis menggunakan relasi dari variabel total fatalities, year, dan probable causes. Sebelumnya, penulis telah mengetahui bahwa terdapat enam kategori penyebab kecelakaan pesawat terbang komersial pada data yang digunakan, yaitu *conflict factors*, *technical factors*, *poor weather*, *human errors*, *crashing the other objects*, dan *disappeared without trace*. Disini penulis akan menganalisis grafik jumlah kecelakaan dan kematian per tahun pada masing-masing kategori tersebut untuk melihat perubahan grafik sebelum dan setelah dibentuknya ICAO dan titik puncak tertinggi dari setiap grafik, baik grafik kecelakaan maupun kematian per tahun.

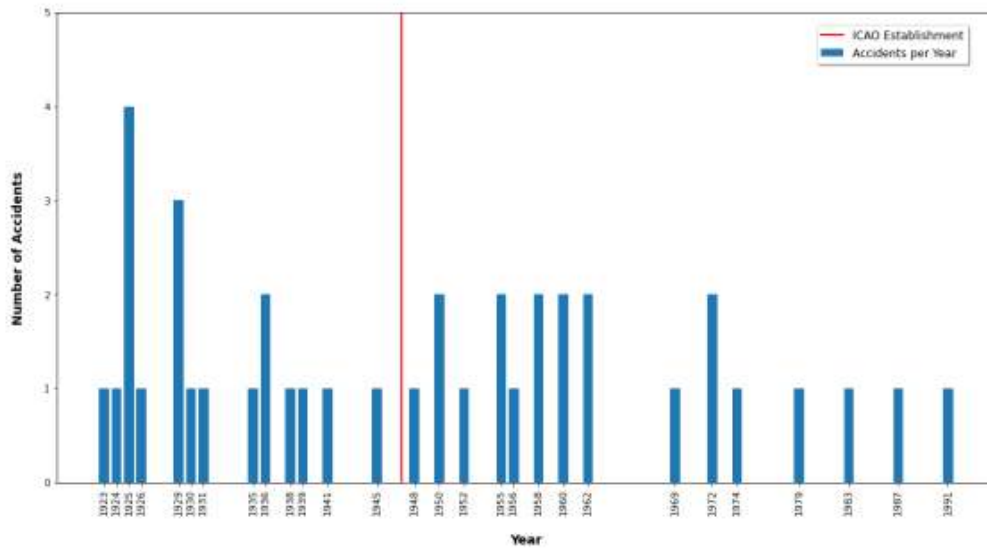
#### **4.2.3.1 Technical Factors**

Kategori penyebab kecelakaan pesawat terbang pertama adalah *technical factors*. Penulis akan menganalisis grafik jumlah kecelakaan dan kematian pesawat terbang komersial dengan penyebab kecelakaan technical factors per tahun, yang

divisualisasikan dengan menggunakan diagram batang.

Visualisasi grafik jumlah kecelakaan per tahun selama periode

tahun 1918-2020 ditunjukkan pada **Gambar 4.21**.



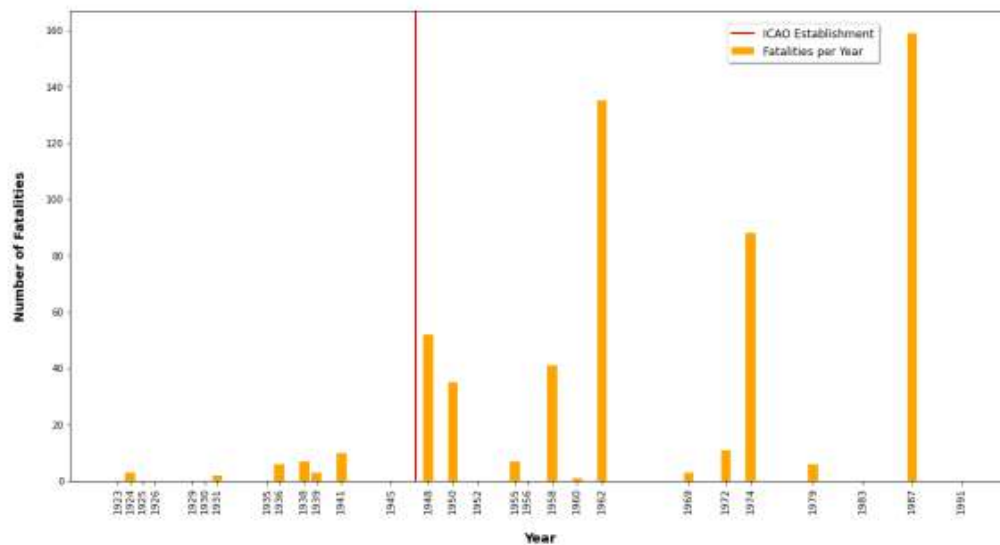
**Gambar 4.21** Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *technical factors*

Berdasarkan diagram diatas, dapat dilihat bahwa selama tahun 1918-2020, telah terjadi 40 kasus kecelakaan pada 28 kategori tahun berbeda. Puncaknya adalah pada tahun 1925 dengan jumlah kecelakaan sebanyak 4 kali dalam satu tahun.

Ketika dilihat lebih jelas, terlihat bahwa grafik jumlah kecelakaan mengalami penurunan yang cukup signifikan yang dimulai dari tahun 1931, dan terus berlanjut hingga tahun 1991. Setelah dibentuk ICAO, grafik jumlah kecelakaan per tahun terlihat tetap stabil, walaupun terjadi kenaikan dan penurunan sebanyak 1 kasus. Dapat diketahui bahwa pada periode tahun 1918-1947 sebelum ICAO dibentuk, telah terjadi kecelakaan sebanyak 19 kali dalam 29 tahun, dengan rata-rata kecelakaan 0.65 per tahun. Sementara pada periode tahun 1948-2020 setelah ICAO dibentuk, kasus kecelakaan yang terjadi sebanyak 21 kali dalam 72 tahun, dengan rata-rata kecelakaan 0.291 per tahun. Dari sini, diketahui bahwa memang pada periode tahun 1948-2020, jumlah kecelakaan lebih banyak dibandingkan periode tahun sebelumnya, tetapi rata-rata kecelakaannya lebih rendah. Selain itu intensitas kecelakaan juga sedikit mengalami penurunan walaupun belum signifikan.



Selanjutnya, penulis akan menganalisis grafik jumlah kematian per tahun yang <sup>4</sup> divisualisasikan dalam bentuk diagram batang pada Gambar 4.22.



**Gambar 4.22** Grafik jumlah kematian pesawat terbang komersial per tahun kategori dengan penyebab kecelakaan *technical factors*

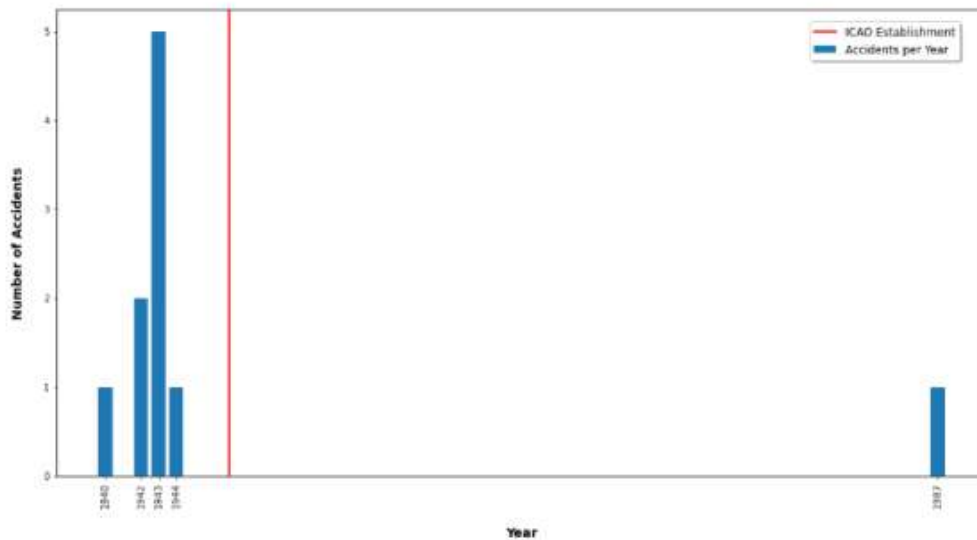
Berdasarkan diagram batang tersebut, dapat dilihat bahwa sama seperti pada analisis grafik jumlah kematian pada jenis aktivitas penerbangan komersial sebelumnya, dimana grafik mengalami peningkatan yang sangat signifikan, terutama



setelah dibentuknya ICAO. Diketahui bahwa dari tahun 1918-2020, total kematian yang dihasilkan adalah sebanyak 569 jiwa, dimana jumlah kematian pada periode tahun 1918-1947 sebanyak 31 jiwa, dan periode tahun 1948-2020 sebanyak 538 jiwa. Dengan jarak yang begitu jauh, cukup masuk akal jika grafik meningkat sangat tajam, walaupun sempat terjadi penurunan jumlah kematian di beberapa tahun. Hal ini menunjukkan bahwa ICAO belum bisa mengatasi jumlah kematian akibat *technical factors* yang terjadi per tahun.

#### 4.2.3.2 Conflict Factors

Setelah menganalisis jumlah kecelakaan per tahun berdasarkan *technical factors*, selanjutnya penulis akan menganalisis grafik jumlah kecelakaan dan kematian pesawat terbang komersial per tahun dengan penyebab kecelakaan *conflict factors*. Visualisasi grafik jumlah kecelakaan per tahun dapat dilihat pada **Gambar 4.23**.



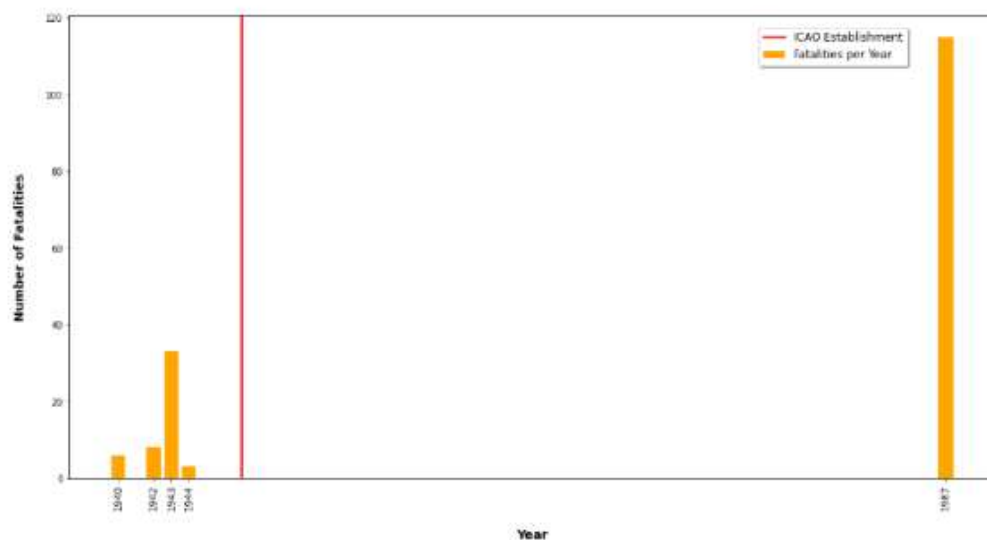
**Gambar 4.23** Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *conflict factors*

Berdasarkan diagram tersebut, dapat dilihat bahwa selama tahun 1918-2020, total kecelakaan yang terjadi sebanyak 10 kasus, dimana 9 diantaranya terjadi pada periode tahun 1918-1947 sebelum dibentuk ICAO, dan 1 kasus terjadi pada periode tahun 1948-2020 setelah dibentuk ICAO. Dengan ini dapat diketahui bahwa grafik jelas mengalami penurunan, dan pembentukan ICAO sukses mengurangi, bahkan menekan jumlah kecelakaan pesawat terbang komersial per tahun dengan

penyebab *conflict factors*. Selain itu, faktor lain yang juga membantu adalah setelah selesainya perang dunia kedua pada tahun 1945, praktis konflik antar setiap negara mulai berkurang, sehingga penerbangan dapat dilakukan dengan lebih aman.

Kemudian, penulis akan menganalisis grafik jumlah kematian per tahun yang divisualisasikan dalam bentuk diagram batang pada **Gambar 4.24**.

Berdasarkan grafik diatas, dapat dilihat bahwa walaupun grafik jumlah kecelakaan per tahu mengalami penurunan yang sangat drastis, tetapi jumlah kematian per tahun yang dihasilkan justru mengalami peningkatan. Cukup menarik dimana pada tahun 1943, dengan jumlah kecelakaan tertinggi sebanyak 5 kasus dalam satu tahun, memiliki jumlah kematian sebanyak 33 jiwa, jauh lebih rendah dibandingkan pada tahun 1987 yang memiliki jumlah kematian sebanyak 115 jiwa, walaupun hanya terjadi 1 kasus kecelakaan saja pada tahun tersebut.

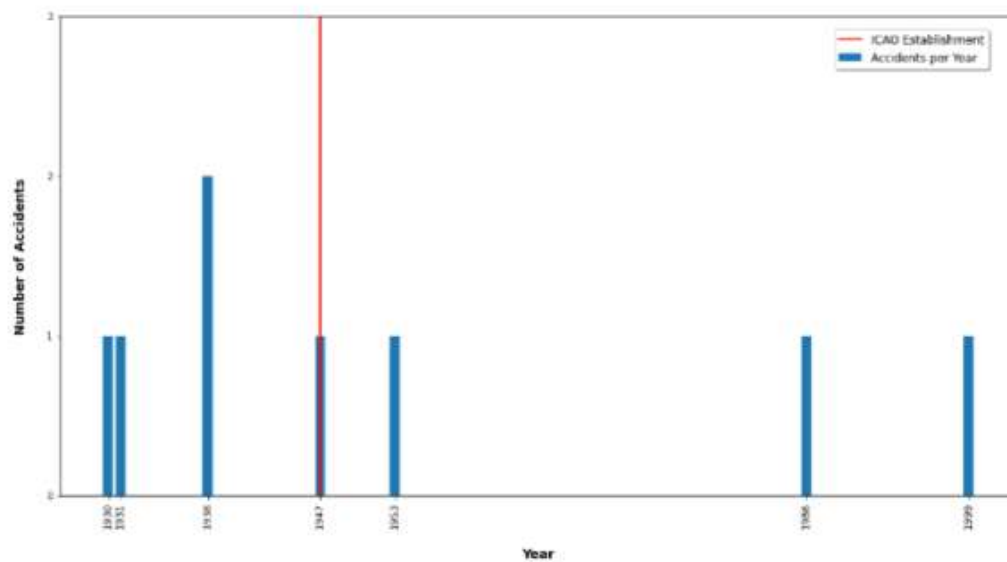


**Gambar 4.24** Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *conflict factors*

#### 4.2.3.3 Human Errors

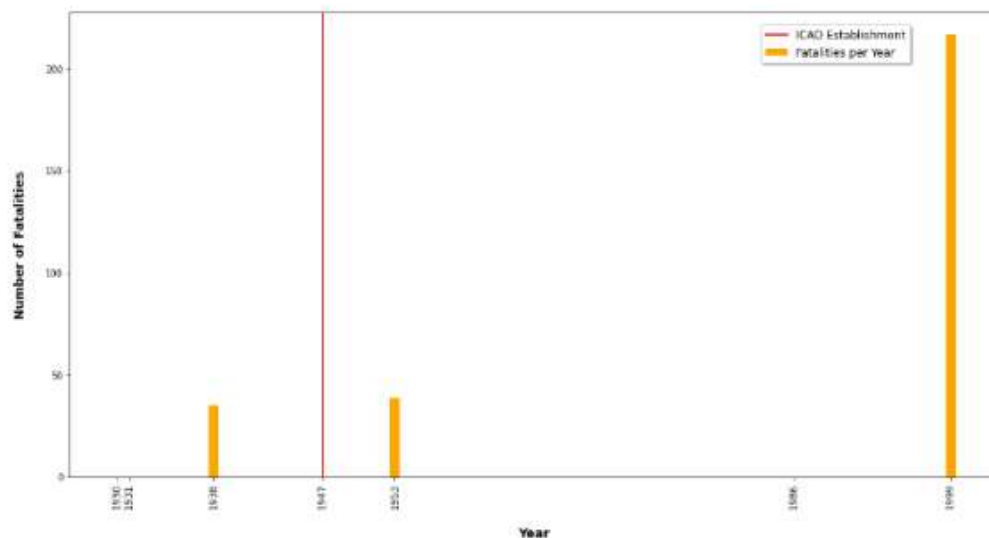
Kategori penyebab kecelakaan yang akan dianalisis selanjutnya adalah *human errors*. Sama seperti pada analisis kategori penyebab kecelakaan sebelumnya, penulis akan menganalisis grafik jumlah kecelakaan dan kematian per tahun selama periode tahun 1918-2020 dengan menggunakan diagram batang. Visualisasi grafik jumlah kecelakaan per tahun dapat dilihat pada **Gambar 4.25**.

Berdasarkan grafik jumlah kecelakaan per tahun diatas, dapat dilihat bahwa 8 kasus kecelakaan pesawat terbang komersial dengan kategori penyebab kecelakaan human errors selama periode tahun 1918-2020. Dari ke 8 kasus kecelakaan tersebut, 5 diantaranya terjadi pada periode tahun 1918-1947, dan 3 lainnya pada periode tahun 1948-2020. Dapat dilihat bahwa grafik mengalami penurunan jumlah kecelakaan per tahun setelah dibentuknya ICAO. Selain itu, intensitas kecelakaan juga mulai berkurang, hal ini dapat dilihat dari gap antar tahun yang jauh antara satu dan lainnya. Rata-rata kecelakaan juga mengalami penurunan, dimana pada periode tahun 1918-1947, rata-ratanya 0.275 per tahun, sementara pada periode tahun 1948-2020, rata-ratanya 0.041 per tahun.



**Gambar 4.25** Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *human errors*

Selanjutnya, penulis akan menganalisis grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan human error yang ditunjukkan pada **Gambar 4.26**.



**Gambar 4.26** Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *human errors*

Berdasarkan diagram diatas, sama seperti analisis sebelumnya, bahwa grafik jumlah kematian selalu mengalami peningkatan setelah dibentuknya ICAO. Dapat dilihat pada grafik bahwa pada periode tahun 1918-1947, menghasilkan total kematian sebanyak 35 jiwa dari 4 kasus kecelakaan, dimana hanya kecelakaan pada kategori tahun 1938 yang menghasilkan jumlah kematian, sedangkan sisanya tidak menghasilkan kematian. Sementara pada periode tahun 1948-2020,

menghasilkan total jumlah kematian sebanyak 256 jiwa dari 3 kasus kecelakaan. Titik puncak pada grafik terjadi pada tahun 1999 yang menghasilkan 239 kematian dari hanya 1 kasus kecelakaan saja, lebih tinggi dibandingkan tahun 1938 yang mengalami 2 kasus kecelakaan dalam satu tahun.

#### 4.2.3.4 Poor Weather

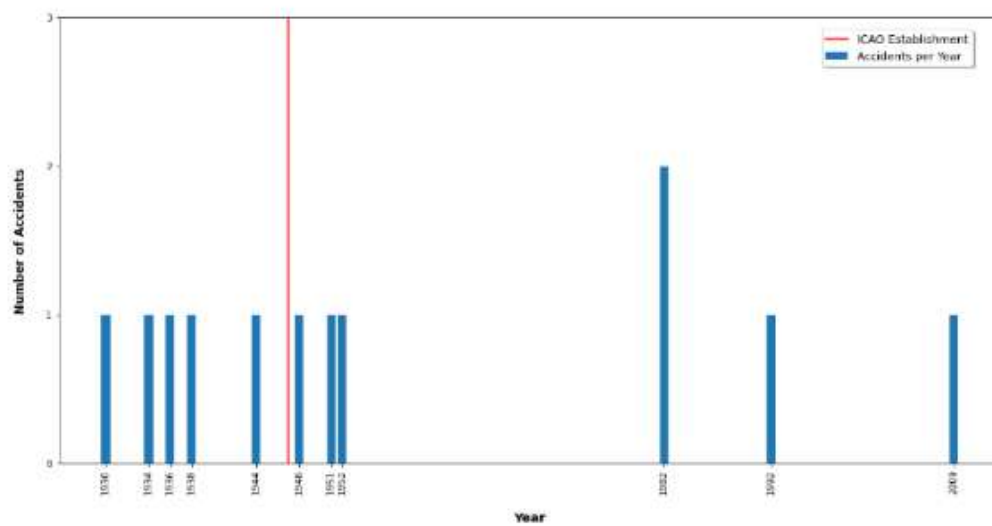
Selanjutnya, penulis akan menganalisis grafik jumlah kecelakaan dan kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan poor weather. Visualisasi grafik jumlah kecelakaan per tahun dapat dilihat pada **Gambar 4.27**.

Berdasarkan grafik tersebut, dapat dilihat bahwa dari tahun 1918-2020, total jumlah kecelakaan yang terjadi sebanyak 12 kali, dengan puncaknya pada tahun 1982 dengan jumlah kecelakaan sebanyak 2 kasus selama satu tahun. Selain itu dapat dilihat bahwa grafik sedikit mengalami peningkatan setelah dibentuknya ICAO, yaitu pada tahun 1982, lalu grafik kembali

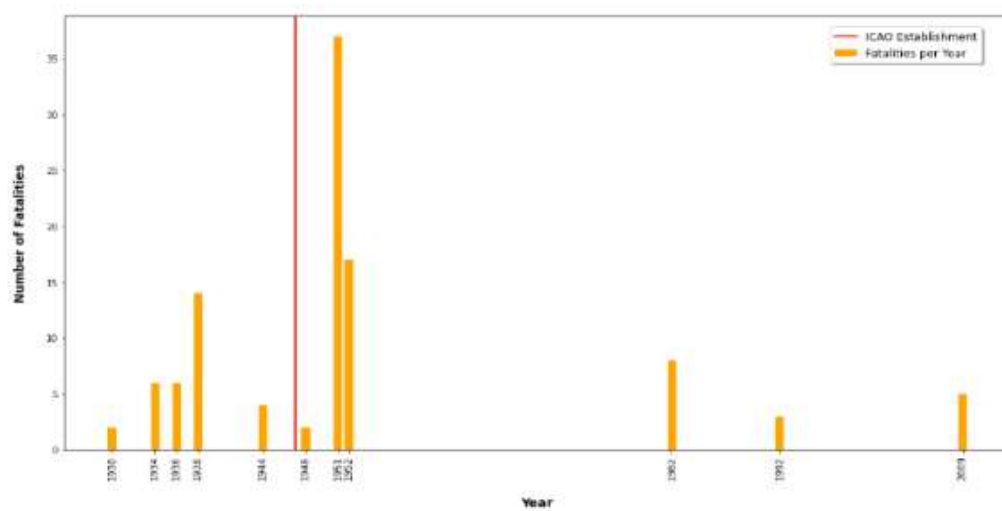


menurun dan stabil. Pada periode tahun 1918-1947 sebelum dibentuknya ICAO, telah terjadi 5 kasus kecelakaan dalam kurun waktu 29 tahun, dengan rata-rata kecelakaan 0.172 per tahun. Sedangkan pada periode tahun 1948-2020 setelah dibentuknya ICAO, kasus kecelakaan yang terjadi sebanyak 7 kali dalam 72 tahun, dengan rata-rata kecelakaan sebesar 0.09 per tahun. Dengan ini, didapatkan bahwa setelah dibentuknya ICAO, jumlah kecelakaan memang mengalami peningkatan, tetapi rata-ratanya mengalami penurunan. Selain itu, intensitas kecelakaan juga mengalami penurunan yang cukup signifikan, yang ditandai dengan jarak antar tahun yang jauh.

Selanjutnya penulis akan menganalisis perubahan grafik jumlah kematian per tahun yang divisualisasikan pada **Gambar 4.28**.



**Gambar 4.27** Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *poor weather*



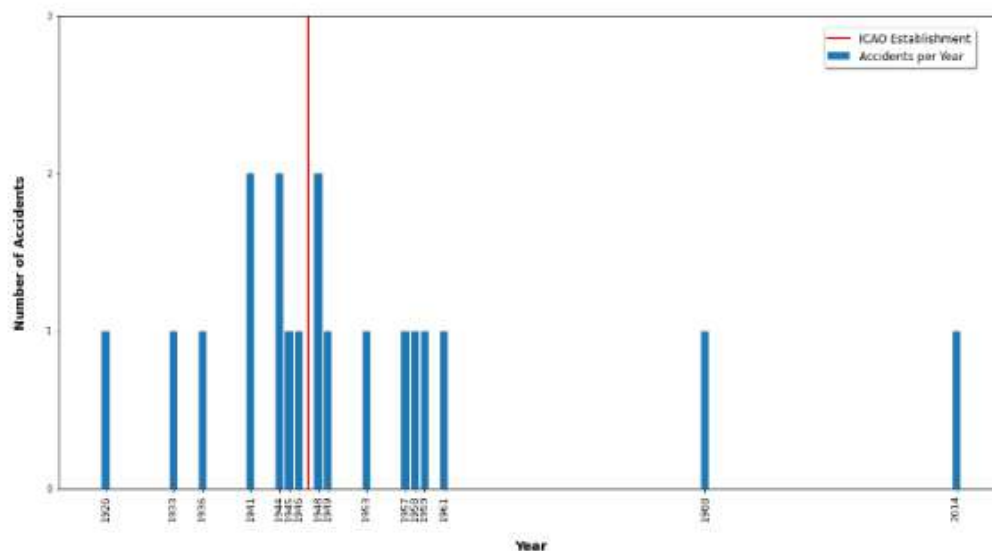
**Gambar 4.28** Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *poor weather*

Berdasarkan diagram batang diatas, dapat dilihat bahwa dari tahun 1918-2020, total jumlah kematian yang dihasilkan oleh kecelakaan pesawat terbang komersial dengan kategori penyebab kecelakaan poor weather sebanyak 104 jiwa, dimana 32 diantaranya dihasilkan pada periode tahun 1918-1947, dan sisanya dihasilkan pada periode tahun 1948-2020. Pada grafik, dapat dilihat bahwa grafik sempat mengalami peningkatan setelah dibentuknya ICAO, dengan puncaknya pada tahun 1951 yang menghasilkan total kematian sebanyak 37 jiwa dalam satu tahun. Walaupun sempat mengalami peningkatan, tetapi grafik kembali mengalami penurunan pada tahun 1982, bahkan dengan jumlah kematian per tahun yang sedikit lebih rendah dibandingkan periode sebelumnya.

#### **4.2.3.5 Disappeared Without Trace**

Kategori penyebab kecelakaan selanjutnya adalah disappeared without trace. Sama seperti sebelumnya, penulis akan menganalisis grafik jumlah kecelakaan dan kematian

pesawat terbang komersial dengan kategori penyebab kecelakaan disappeared without trace per tahun selama periode tahun 1918-2020. Visualisasi grafik jumlah kecelakaan per tahun dapat dilihat pada **Gambar 4.29**.

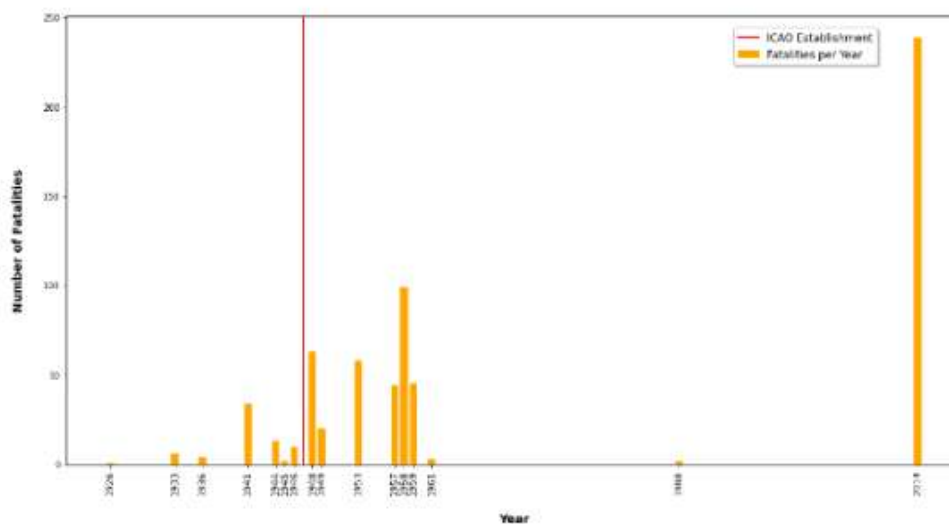


**Gambar 4.29** Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *disappeared without trace*

Berdasarkan diagram batang diatas, diketahui bahwa dari tahun 1918-2020, terdapat 19 kasus kecelakaan, dimana 9 diantaranya terjadi pada periode tahun 1918-1947, dan 10

sisanya terjadi pada periode tahun 1948-2020. Untuk perubahan grafik, tidak terlihat adanya peningkatan atau penurunan yang signifikan antara periode tahun sebelum dan setelah dibentuknya ICAO, atau dalam kata lain grafik cukup stabil. Untuk rata-rata kecelakaan per tahun, pada periode tahun 1918-1947 memiliki rata-rata kecelakaan sebesar 0.62 per tahun, sedangkan pada periode tahun 1948-2020 memiliki rata-rata kecelakaan sebesar 0.138 per tahun. Sehingga walaupun jumlah kecelakaan pada periode tahun 1948-2020 sedikit lebih banyak, tetapi rata-rata kecelakaannya lebih rendah. Selain itu juga intensitas kecelakaan per tahun sudah mulai mengalami penurunan yang terlihat dimulai dari tahun 1961.

Selanjutnya untuk grafik jumlah kematian akibat pesawat terbang komersial dengan kategori penyebab kecelakaan disappeared without trace <sup>4</sup> divisualisasikan dalam bentuk diagram batang pada **Gambar 4.30**.



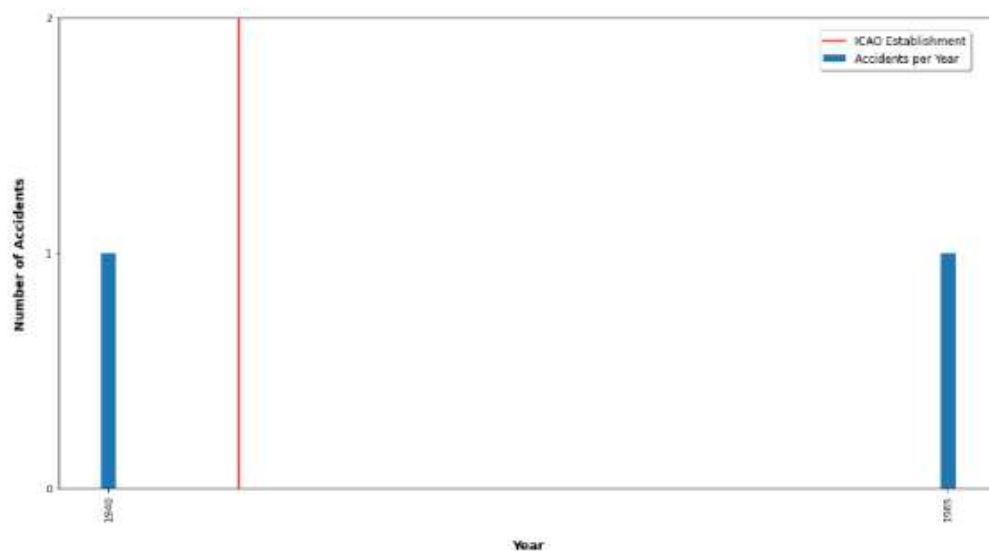
**Gambar 4.30** Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *disappeared without trace*

Berdasarkan diagram diatas, dapat dilihat bahwa dari tahun 1918-2020, total kematian yang dihasilkan adalah sebanyak 643 jiwa, dimana 70 diantaranya dihasilkan dari periode tahun 1918-1947, sementara 543 lainnya dihasilkan dari periode tahun 1948-2020. Selain itu, dapat dilihat juga bahwa grafik mengalami peningkatan pada periode tahun 1948-2020, dengan puncaknya pada tahun 2014 yang menghasilkan jumlah kematian sebanyak 239 jiwa, walaupun hanya dari 1 kasus

kecelakaan saja. Sehingga didapatkan bahwa walaupun setelah dibentuknya ICAO, tetapi grafik jumlah kematian per tahun tidak mengalami penurunan, bahkan justru mengalami peningkatan.

#### **4.2.3.6 Crashing The Other Objects**

Kategori penyebab kecelakaan yang terakhir adalah crashing the other objects. Penulis juga akan menganalisis perubahan grafik jumlah kecelakaan dan kematian pesawat terbang komersial dengan kategori penyebab kecelakaan ini per tahun selama periode tahun 1918-2020. Masing-masing grafik akan divisualisasikan dalam bentuk diagram batang. Visualisasi grafik jumlah kecelakaan per tahun dapat dilihat pada **Gambar 4.31**.



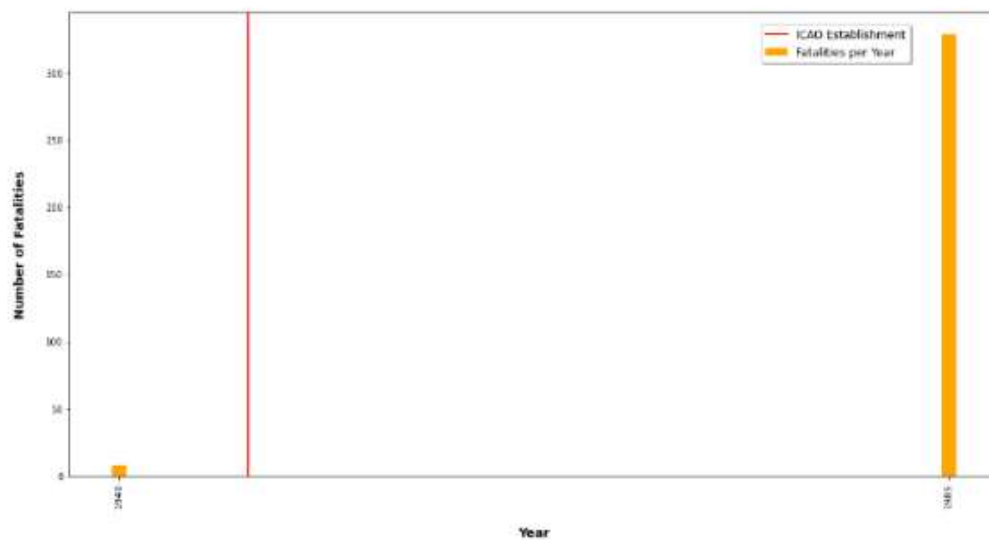
**Gambar 4.31** Grafik jumlah kecelakaan pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *crashing the other objects*

Berdasarkan diagram diatas, dapat dilihat bahwa dari tahun 1918-2020, hanya terjadi 2 kasus kecelakaan pesawat terbang komersial dengan kategori penyebab kecelakaan *crashing the other objects*, yaitu pada tahun 1940 dan 1985. Karena hanya terjadi 2 kecelakaan pada 2 kategori tahun yang berbeda, maka grafik tidak mengalami peningkatan ataupun penurunan. Untuk rata-rata kecelakaan per tahun, pada periode tahun 1918-1947 memiliki rata-rata 0.034 per tahun, sedangkan



untuk periode tahun 1948-2020 memiliki rata-rata 0.013 per tahun. Sehingga, walaupun kedua periode tersebut sama-sama memiliki 1 kasus kecelakaan saja, tetapi rata-rata kecelakaan pada periode tahun 1948-2020 setelah dibentuknya ICAO lebih rendah.

Selanjutnya, penulis akan menganalisis grafik jumlah kematian per tahun yang <sup>4</sup> divisualisasikan dalam bentuk diagram batang seperti pada Gambar 4.32.



**Gambar 4.32** Grafik jumlah kematian pesawat terbang komersial per tahun dengan kategori penyebab kecelakaan *crashing the other objects*

Berdasarkan diagram batang diatas, dapat dilihat bahwa dari tahun 1918-2020, total jumlah kematian yang dihasilkan adalah sebanyak 337 jiwa, dimana 8 diantaranya dihasilkan dari kasus kecelakaan pada periode tahun 1918-1947, dan 329 sisanya dihasilkan dari kasus kecelakaan pada periode tahun 1948-2020. Terlihat grafik mengalami peningkatan jumlah kematian per tahun yang sangat signifikan, walaupun hanya terjadi 1 kasus kecelakaan saja pada kedua tahun tersebut. Hal ini sama dengan analisis pada kategori penyebab kecelakaan sebelumnya, dimana pada periode tahun 1948-2020, walaupun hanya memiliki 1-2 kasus kecelakaan dalam satu tahun, tetapi jumlah kematian yang dihasilkan sangat banyak, bahkan lebih banyak dibandingkan kategori tahun dengan jumlah kecelakaan tertinggi pada periode tahun 1918-1947.

#### **4.2.4 Summary Analysis**

Setelah menganalisis data kecelakaan pesawat terbang komersial BAAA dengan menggunakan relasi antar variabel

yang digunakan, penulis mendapati bahwa hampir seluruh grafik jumlah kecelakaan per tahun mengalami penurunan yang cukup drastis pada periode tahun 1948-2020, setelah dibentuknya ICAO. Walaupun terdapat beberapa grafik jumlah kecelakaan yang sedikit mengalami peningkatan, tetapi rata-rata dan intensitas kecelakaan per tahunnya tetap lebih rendah dibandingkan periode tahun 1918-1947 sebelum dibentuknya ICAO. Walaupun mengalami penurunan grafik jumlah kecelakaan per tahun, tetapi grafik jumlah kematian justru meningkat dengan tajam. Contohnya seperti pada tahun 1985, walaupun hanya mengalami satu kecelakaan dalam satu tahun, tetapi jumlah kematian yang dihasilkan sebesar 329 jiwa, jauh lebih banyak dibandingkan tahun 1943 dengan jumlah kasus kecelakaan per tahun tertinggi pada data yang hanya menghasilkan jumlah kematian sebanyak 33 jiwa.

Untuk aktivitas penerbangan scheduled revenue flight, grafik jumlah kecelakaan jelas mengalami penurunan setelah periode tahun 1948-2020. Walaupun begitu, tetapi grafik jumlah

kematian justru meningkat dengan signifikan, walaupun dengan jumlah kecelakaan yang lebih sedikit. <sup>2</sup> Berdasarkan Gambar 4.14, dapat dilihat bahwa pada periode tahun 1918-1947 sebelum dibentuknya ICAO, jumlah kematian per tahun berada pada angka 50 kebawah, sementara pada periode tahun 1948-2020 setelah dibentuknya ICAO, jumlah kematian per tahun meningkat tajam, seperti tahun 1985, 1987, 1999, dan 2014 dengan jumlah kematian diatas 200 jiwa, walaupun dengan hanya 1-2 kasus kecelakaan saja. Hal yang sama juga terjadi pada aktivitas penerbangan charter / taxi flight dan cargo flight, dimana kenaikan grafik jumlah kematian per tahun justru dibarengi dengan grafik jumlah kecelakaan yang stabil dan tidak mengalami penurunan. Lain halnya dengan kategori aktivitas penerbangan postal flight, dimana kecelakaan hanya terjadi pada periode tahun 1918-1947 saja, sehingga tidak ada satu pun kasus kecelakaan pesawat terbang komersial pada periode tahun 1948-2020. Selain itu, dapat dilihat juga bahwa jumlah kematian pada aktivitas penerbangan komersial cargo flight dan postal flight

memiliki range yang sangat jauh dibandingkan penerbangan scheduled revenue flight dan charter / taxi. Hal ini dikarenakan aktivitas penerbangan cargo dan postal hanya bertujuan untuk mengantarkan barang dan surat saja.

Dari keseluruhan penjelasan diatas, didapatkan bahwa grafik jumlah kecelakaan mengalami penurunan seiring berkembangnya tahun, tetapi hal tersebut justru berbanding terbalik dengan grafik jumlah kematian per tahun yang justru mengalami peningkatan yang sangat tajam. Fenomena tersebut terjadi seiring evolusi dari pesawat terbang itu sendiri. Semakin bertambahnya tahun, maka kebutuhan masyarakat mengenai transportasi udara juga semakin tinggi per tahunnya, yang menyebabkan setiap industri penerbangan di seluruh dunia terus melakukan suatu perubahan untuk memenuhi permintaan masyarakat. Selain itu, penurunan grafik jumlah kecelakaan pada periode tahun 1948-2020 setelah dibentuknya ICAO dapat terjadi karena adanya peningkatan dan inovasi terhadap sistem keselamatan, teknologi, dan design pada pesawat terbang yang

dilakukan oleh setiap industri penerbangan. Seperti yang diketahui bahwa ICAO memiliki tugas untuk menstandarisasi aturan dan ketentuan keselamatan penerbangan, sehingga dapat menjamin keselamatan penerbangan dan pertumbuhan penerbangan sipil yang teratur diseluruh dunia. Kemudian grafik peningkatan jumlah kematian per tahun juga terjadi karena adanya evolusi dari ukuran suatu pesawat terbang. Sejak tahun 1960, perjalanan internasional menjadi semakin populer, sehingga volume penumpang semakin meningkat secara signifikan. Oleh karena itu, pabrikan-pabrikan pesawat terbang di seluruh dunia melakukan evolusi terhadap ukuran pesawat mereka untuk dapat menampung volume penumpang yang lebih banyak. Hingga pada tahun 1969, dibangun sebuah pesawat jet jumbo oleh Boeing, dengan nama Boeing 747. Pesawat jet jumbo tersebut tentunya dapat mengangkut lebih banyak penumpang dibandingkan pesawat jet biasa. Walaupun memiliki keuntungan dengan bisa mengangkut penumpang yang lebih banyak, tetapi tentunya terdapat konsekuensi dari evolusi

tersebut. Salah satu konsekuensinya adalah ketika terjadi sebuah kecelakaan yang melibatkan pesawat jet jumbo tersebut, maka jumlah kematian yang dihasilkan juga akan lebih banyak dibandingkan pesawat jet biasa. Hal ini dikarenakan jumlah penumpang yang diangkut juga lebih banyak dibandingkan pesawat jet biasa. Pada data yang digunakan, hal tersebut sesuai seperti pada tahun 1985, yang hanya mengalami satu kasus kecelakaan saja, tetapi jumlah kematian yang dihasilkan jauh lebih banyak dibandingkan kategori tahun lainnya dengan jumlah kecelakaan per tahun yang lebih banyak.

14

*This Page Intentionally Left Blank*



# CHAPTER 5 CONCLUSIONS



*This Page Intentionally Left Blank*

## CHAPTER 5

### CONCLUSIONS

Penulisan yang dilakukan bertujuan untuk mendeskripsikan dan menganalisis relasi-relasi setiap variabel pada data kecelakaan pesawat terbang komersial BAAA untuk melihat apakah terjadi perubahan terhadap grafik kecelakaan dan kematian pesawat terbang komersial per tahun setelah dibentuknya ICAO, dengan menggunakan metode *descriptive statistics* dan EDA. Berdasarkan hasil penulisan, didapatkan bahwa setelah dibentuknya ICAO sebagai organisasi keselamatan penerbangan sipil internasional, terjadi penurunan yang signifikan terhadap grafik jumlah dan intensitas kecelakaan pesawat terbang komersial per tahun. Walaupun grafik jumlah kecelakaan per tahun mengalami penurunan, tetapi hal tersebut berbanding terbalik dengan grafik jumlah kematian per tahun,

dimana justru grafik jumlah kematian per tahun mengalami peningkatan yang juga signifikan setelah pada periode tahun 1948-2020 setelah dibentuknya ICAO. <sup>43</sup> Salah satu faktor yang mendukung fenomena ini adalah evolusi ukuran pesawat yang dilakukan industri penerbangan untuk memenuhi permintaan masyarakat.

Kedua metode yang digunakan cukup efektif untuk mendeskripsikan dan menganalisis relasi dari setiap variabel pada data yang digunakan. Dengan metode descriptive statistics, penulis dapat mengetahui jenis distribusi data, titik pusat distribusi data, persebaran setiap titik data dari titik pusat distribusi data, dan menentukan jenis kurva distribusi data. Sementara dengan metode EDA, penulis dapat menganalisis relasi dari setiap variabel yang digunakan. Kedua metode tersebut digunakan dengan bantuan perhitungan statistik dan visualisasi dengan berbagai jenis diagram. Hal tersebut jelas lebih baik dibandingkan metode sebelumnya yang masih menggunakan pemakaran dan pengalaman saja.

Kesimpulan akhir dari penulisan yang dilakukan adalah, pembentukan ICAO sebagai organisasi keselamatan penerbagan sipil internasional berhasil mengurangi jumlah dan intensitas kecelakaan pesawat terbang komersial yang terjadi dengan sangat signifikan. Tetapi ICAO tidak berhasil mengurangi grafik jumlah kematian per tahun. Hal ini dikarenakan seiring berkembangnya tahun, maka akan selalu terjadi evolusi pada ukuran suatu pesawat komersial terbang untuk memenuhi permintaan masyarakat yang semakin meningkat setiap tahunnya. Sehingga dari seluruh hasil penulisan yang dilakukan, didapatkan bahwa hipotesis awal yang mengatakan bahwa grafik jumlah kecelakaan dan kematian pesawat terbang komersial mengalami penurunan setelah dibentuknya ICAO terbukti salah, karena hanya grafik jumlah kecelakaan saja yang mengalami penurunan.

*This Page Intentionally Left Blank*



A twin-engine turboprop aircraft is flying in a blue sky with scattered white clouds. Below the aircraft, a road with white lane markings is visible. A silver car is driving on the road. In the background, there are palm trees and a building. A white fence runs along the road, with a blue sign featuring a portrait of a man. The text "CHAPTER 6 CLOSING STATEMENTS" is overlaid on the image in white, bold, sans-serif font.

## CHAPTER 6 CLOSING STATEMENTS

*This Page Intentionally Left Blank*



## CHAPTER 6

### CLOSING STATEMENT

Demikian tulisan yang dapat kami paparkan dalam buku ini. Tentunya masih banyak kekurangan dan kelemahan dalam pembuatan buku ini karena keterbatasan pengetahuan dari penulis. Maka dari itu kami menyadari bahwa buku ini masih jauh dari kata sempurna, dan kedepannya penulis akan lebih memfokuskan dalam menjelaskan mengenai tulisan diatas dengan sumber yang lebih banyak, dan tentunya dapat dipertanggungjawabkan. kami harap para pembaca yang akan membaca buku ini dapat memberikan kritik dan saran yang membangun, sehingga kami dapat menyempurnakan buku ini pada kesempatan berikutnya.

Akhir kata semoga buku ini semua dapat bermanfaat bagi kami dan bagi para pembacanya, dan menjadi amal jariyah bagi kita semua dalam berbagi ilmu ini. Terimakasih atas kesempatannya.

## GLOSARIUM

ICAO ( <i>International Civil Aviation Organization</i> )	3
EDA ( <i>Exploratory Data Analysis</i> )	4
BAAA ( <i>Bureau of Aircraft Accidents Archives</i> )	5
<sup>110</sup> CFIT ( <i>Controlled Flight Into Terrain</i> )	10
HFACS ( <i>Human Factors Analysis and Classification</i> )	10
<sup>47</sup> EGPWS ( <i>Enhanced Ground Proximity Warning System</i> )	10
GPWS ( <i>Ground Proximity Warning System</i> )	10
LGA ( <i>Longitudinal Acceleration</i> )	12
VTA ( <i>Vertical Acceleration</i> )	12
LTA ( <i>Lateral Acceleration</i> )	12
PTCH ( <i>Pitch Angle</i> )	12
<sup>120</sup> PCA ( <i>Principal Component Analysis</i> )	13
LDA ( <i>Linear Discriminant Analysis</i> )	13
SARPs ( <i>Standards and recommended practices</i> )	14
IQR ( <i>Interquartile Range</i> )	25

*This Page Intentionally Left Blank*

# INDEX

---

## **A**

*Aviation* · iii, 1, 21, 22, 24, 26, 27, 172, 181, 184

---

## **B**

*Bureau of Aircraft Accidents Archives* · 6, 172, 191

---

## **D**

Data Cleaning · 32

*data collection* · 29, 61, 63, 185

Data Integration · 34

*data preprocessing* · 29, 31, 32, 61, 64, 69, 81, 186

*data qualitative* · 36, 42, 54, 76, 104, 108

*data quantitative* · 36, 41, 70, 76, 81, 104

Data Reduction · 32

Data Transformation · 33

*descriptive statistics* · iii, iv, 4, 6, 7, 34, 61, 69, 77, 81, 112, 165, 166, 182, 186,  
187

*determine analysis objectives* · 29, 61

---

## **E**

Exploratory Data Analysis · 8, 55, 77, 112, 172, 190

---

## *I*

*ICAO* · i, iii, 1, 4, 5, 6, 8, 20, 21, 22, 24, 25, 26, 63, 67, 78, 105, 113, 114, 116, 119, 120, 122, 124, 126, 128, 130, 132, 133, 135, 137, 139, 141, 144, 146, 147, 149, 151, 153, 155, 157, 158, 160, 165, 167, 172, 181, 183, 184  
*IQR* · 44, 46, 54, 73, 74, 93, 94, 95, 97, 98, 99, 173

---

## *K*

keselamatan · 5, 6, 18, 20, 22, 63, 127, 160, 165, 167  
keselamatan penerbangan · 20, 63, 161

---

## *M*

*mean* · 37, 38, 39, 40, 41, 47, 49, 50, 54, 71, 72, 73, 86, 87, 88, 89, 90, 91, 92, 182, 187, 192  
Measures of Central Tendency · 37, 71, 86, 88, 91, 191  
Measures of Dispersion · 42, 72, 93, 94, 97  
Measures of Frequency · 37, 71, 82  
*median* · 37, 38, 39, 40, 41, 46, 71, 72, 86, 87, 88, 89, 90, 91, 92, 93, 94, 192

---

## *N*

non-scheduled · 67, 123

---

## *O*

*outliers* · 32, 39, 40, 41, 43, 45, 46, 50, 52, 53, 55, 72, 74, 76, 90, 92, 94, 96, 98, 99, 102, 104

---

## *P*

penerbangan · 1, 2, 3, 5, 6, 12, 14, 17, 20, 21, 22, 23, 24, 25, 26, 27, 63, 66, 67, 69, 104, 105, 106, 107, 108, 118, 119, 120, 122, 123, 124, 125, 127, 128, 129, 130, 131, 133, 135, 139, 142, 158, 160, 165

penerbangan komersial · 1, 26, 105, 106, 107, 108, 118, 119, 123, 127, 130, 139, 159

Penerbangan sipil · 24

---

## *R*

*range* · 44, 45, 51, 54, 73, 75, 93, 94, 99, 159, 188

---

## *S*

*Scheduled Revenue Flight* · 106, 108, 119

Skewness and Kurtosis · 99, 189

*standard deviation* · 44, 49, 54, 73, 94, 189

---

## *V*

Variables Selection · 68

*variance* · 44, 46, 47, 48, 50, 54, 73, 93, 94, 188

*This Page Intentionally Left Blank*



## BIBLIOGRAPHY

- [1] J. Huang, “Aviation safety, ICAO and obligations erga omnes,” *Chinese J. Int. Law*, vol. 8, no. 1, pp. 63–79, 2009, doi: 10.1093/chinesejil/jmn039.
- [2] T. G. Puranik, N. Rodriguez, and D. N. Mavris, “Towards online prediction of safety-critical landing metrics in aviation using supervised machine learning,” *Transp. Res. Part C Emerg. Technol.*, vol. 120, no. April, p. 102819, 2020, doi: 10.1016/j.trc.2020.102819.
- [3] V. Peysakhovich, O. Lefrançois, F. Dehais, and M. Causse, “The neuroergonomics of aircraft cockpits: The four stages of eye-tracking integration to enhance flight safety,” *Safety*, vol. 4, no. 1, 2018, doi: 10.3390/safety4010008.
- [4] “Transportation Fatalities by Mode | Bureau of Transportation Statistics.” <https://www.bts.gov/content/transportation-fatalities-mode> (accessed Jun. 16, 2021).
- [5] D. Kelly and M. Efthymiou, “An analysis of human factors in fifty controlled flight into terrain aviation accidents from 2007 to 2017,” *J. Safety Res.*, vol. 69, pp. 155–165, 2019, doi: 10.1016/j.jsr.2019.03.009.
- [6] R. M. A. Valdés *et al.*, “Development of safety performance functions (SPFs) to analyse and predict aircraft loss of separation in accordance with the characteristics of the airspace,” *Reliab. Eng. Syst. Saf.*, vol. 186, no. January, pp. 143–161, 2019, doi:

- 10.1016/j.ress.2019.02.007.
- [7] P. Illankoon, P. Tretten, and U. Kumar, “A prospective study of maintenance deviations using HFACS-ME,” *J. Ind. Ergon.*, vol. 74, no. July, p. 102852, 2019, doi: 10.1016/j.ergon.2019.102852.
- [8] J. Oehling and D. J. Barry, “Using machine learning methods in airline flight data monitoring to generate new operational safety knowledge from existing data,” *Saf. Sci.*, vol. 114, no. December 2018, pp. 89–104, 2019, doi: 10.1016/j.ssci.2018.12.018.
- [9] K. W. Herbst *et al.*, “Don’t be mean, be above average: Understanding data distribution and descriptive statistics,” *J. Pediatr. Urol.*, vol. 16, no. 5, p. 712, Oct. 2020, doi: 10.1016/j.jpuro.2020.09.007.
- [10] H. B. Sankaranarayanan, G. Agarwal, and V. Rathod, “An exploratory data analysis of airport wait times using big data visualisation techniques,” *2016 Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solut. CSITSS 2016*, pp. 324–329, 2016, doi: 10.1109/CSITSS.2016.7779379.
- [11] H. Lee, G. Li, A. Rai, and A. Chattopadhyay, “Real-time anomaly detection framework using a support vector regression for the safety monitoring of commercial aircraft,” *Adv. Eng. Informatics*, vol. 44, no. February, p. 101071, 2020, doi: 10.1016/j.aei.2020.101071.
- [12] K. Iwadare and T. Oyama, “Statistical Data Analyses on Aircraft Accidents in Japan: Occurrences, Causes and Countermeasures,” *Am. J. Oper. Res.*, vol. 05, no. 03, pp. 222–245, 2015, doi: 10.4236/ajor.2015.53018.
- [13] T. Makaba, W. Doorsamy, and B. S. Paul, “Exploratory framework for analysing road traffic accident data with variation on Gauteng province data,” *Cogent Eng.*, vol. 7, no. 1, 2020, doi: 10.1080/23311916.2020.1834659.

- [14] <sup>59</sup> K. Button, A. Clarke, G. Palubinskas, R. Stough, and M. Thibault, “Conforming with ICAO safety oversight standards,” *J. Air Transp. Manag.*, vol. 10, no. 4, pp. 249–255, 2004, doi: 10.1016/j.jairtraman.2004.02.004. <sup>61</sup>
- [15] “What is The ICAO, and What Do They Do?” <https://aerocorner.com/blog/icao/> (accessed Jun. 10, 2021).
- [16] P. Dumberry, “Annexes 1 to 3,” *State Succession to Int. Responsib.*, pp. 417–439, 2009, doi: 10.1163/ej.9789004158825.i-520.46.
- [17] I. Standards, R. Practices, and I. C. Aviation, *Aircraft Accident and Incident Investigation Catalogue of ICAO Publications*, no. July. 2001.
- [18] “airplane - Civil aircraft | Britannica.” <https://www.britannica.com/technology/airplane/Civil-aircraft> (accessed Jun. 13, 2021).
- [19] <sup>22</sup> ICAO, “Review of the Classification and Definitions Used for Civil Aviation Activities,” no. August 2008, pp. 1–9, 2009.
- [20] ICAO, *Operation of Aircraft*, no. July. 2010.
- [21] <sup>67</sup> “5 Steps of the Data Analysis Process.” <https://learn.g2.com/data-analysis-process> (accessed Jun. 13, 2021).
- [22] <sup>23</sup> A. Famili, W. M. Shen, R. Weber, and E. Simoudis, “Data preprocessing and intelligent data analysis,” *Intell. Data Anal.*, vol. 1, no. 1, pp. 3–23, 1997, doi: 10.3233/IDA-1997-1102.
- [23] <sup>36</sup> “The 5 Steps of the Data Analysis Process | by Kunal Gohrani | Medium.” [https://medium.com/@kunal\\_gohrani/the-5-steps-of-the-data-analysis-process-2512ba6ac31e](https://medium.com/@kunal_gohrani/the-5-steps-of-the-data-analysis-process-2512ba6ac31e) (accessed Jun. 15,

- 2021).
- [24] E. Paradis, B. O'Brien, L. Nimmon, G. Bandiera, and M. A. T. Martimianakis, "Design: Selection of Data Collection Methods," *J. Grad. Med. Educ.*, vol. 8, no. 2, pp. 263–264, 2016, doi: 10.4300/JGME-D-16-00098.1.
- [25] V. O. Ajayi, "Primary Sources of Data and Secondary Sources of Data," no. September, pp. 1–6, 2017, doi: 10.13140/RG.2.2.24292.68481.
- [26] S. Lefever, M. Dal, and Á. Matthíasdóttir, "Online data collection in academic research: Advantages and limitations," *Br. Educ. Technol.*, vol. 38, no. 4, pp. 574–582, 2007, doi: 10.1111/j.1467-8535.2006.00638.x.
- [27] D. Barapatre and A. Vijayalakshmi, "Data preparation on large datasets for data science," *Asian J. Pharm. Clin. Res.*, vol. 10, pp. 485–488, 2017, doi: 10.22159/ajpcr.2017.v10s1.20526.
- [28] H. Benhar, A. Idri, and J. L. Fernández-Alemán, "Data preprocessing for heart disease classification: A systematic literature review," *Comput. Methods Programs Biomed.*, vol. 195, p. 105635, Oct. 2020, doi: 10.1016/j.cmpb.2020.105635.
- [29] S. García, J. Luengo, and F. Herrera, "Tutorial on practical tips of the most influential data preprocessing algorithms in data mining," *Knowledge-Based Syst.*, vol. 98, pp. 1–29, 2016, doi: 10.1016/j.knosys.2015.12.006.
- [30] J. J. Davis and A. J. Clark, "Data preprocessing for anomaly based network intrusion detection: A review," *Comput. Secur.*, vol. 30, no. 6–7, pp. 353–375, 2011, doi: 10.1016/j.cose.2011.05.008.
- [31] "statistics | Definition, Types, & Importance | Britannica." <https://www.britannica.com/science/statistics> (accessed Jun. 11, 2021).



- [32] <sup>49</sup> M. J. Fisher and A. P. Marshall, “Understanding descriptive statistics,” *Aust. Crit. Care*, vol. 22, no. 2, pp. 93–97, 2009, doi: 10.1016/j.aucc.2008.11.003.
- [33] <sup>74</sup> “Intro to Descriptive Statistics. Descriptive Statistical Analysis helps...”<sup>95</sup> by Niklas Donges | Towards Data Science.” <https://towardsdatascience.com/intro-to-descriptive-statistics-252e9c464ac9> (accessed Jun. 11, 2021).
- [34] <sup>51</sup> J. R. Dettori and D. C. Norvell, “The Anatomy of Data,” *Glob. Spine J.*, vol. 8, no. 3, pp. 311–313, 2018, doi: 10.1177/2192568217746998.
- [35] <sup>40</sup> G. Marshall and L. Jonker, “An introduction to descriptive statistics: A review and practical guide,” *Radiography*, vol. 16, no. 4, pp. e1–e7, 2010, doi: 10.1016/j.radi.2010.01.001.
- [36] E. Allaj, “Measuring variability and association for categorical data,” *Fuzzy Sets Syst.*, vol. 1, pp. 1–15, 2020, doi: 10.1016/j.fss.2020.11.018.
- [37] P. C. O’Brien and M. A. Shampo, “Descriptive statistics,” *Mayo Clinic Proceedings*, 1981. <https://conjointly.com/kb/descriptive-statistics/> (accessed Jun. 10, 2021).
- [38] <sup>10</sup> S. S. Durduran, “Automatic classification of high resolution land cover using a new data weighting procedure: The combination of k-means clustering algorithm and central tendency measures (KMC-CTM),” *Appl. Soft Comput. J.*, vol. 35, pp. 136–150, 2015, doi: 10.1016/j.asoc.2015.06.025.
- [39] <sup>8</sup> K. A. Adamson and S. Prion, “Making sense of methods and measurement: Measures of central tendency,” <sup>116</sup> *Simul. Nurs.*, vol. 9, no. 12, pp. e617–e618, 2013, doi: 10.1016/j.ecns.2013.04.003.

- [40] S. Manikandan, “Measures of central tendency: The mean,” *J. Pharmacol. Pharmacother.*, vol. 2, no. 2, pp. 140–142, 2011, doi: 10.4103/0976-500X.81920.
- [41] S. Manikandan, “Measures of central tendency: Median and mode,” vol. 2, no. 3, pp. 214–215, 2011, doi: 10.4103/0976-500X.83300.
- [42] A. Kołacz and P. Grzegorzewski, “Measures of dispersion for multidimensional data,” *Eur. J. Oper. Res.*, vol. 251, no. 3, pp. 930–937, 2016, doi: 10.1016/j.ejor.2016.01.011.
- [43] S. Manikandan, “Measures of dispersion,” *J. Pharmacol. Pharmacother.*, vol. 2, no. 4, pp. 315–316, 2011, doi: 10.4103/0976-500X.85931.
- [44] “Measures of Variability: Range, Interquartile Range, Variance, and Standard Deviation - Statistics By Jim,” 2018. <https://statisticsbyjim.com/basics/variability-range-interquartile-variance-standard-deviation/> (accessed Jun. 11, 2021).
- [45] K. A. Adamson and S. Prion, “Making sense of methods and measurement: Measures of variability,” *Clin. Sim Nurs.*, vol. 9, no. 11, pp. e559–e560, 2013, doi: 10.1016/j.ecns.2013.04.002.
- [46] S. Ahmad, “On Efficient Monitoring of Process Dispersion using Interquartile Range,” *Open J. Appl. Sci.*, vol. 02, no. 04, pp. 39–43, 2012, doi: 10.4236/ojapps.2012.24b010.
- [47] C. Vanlalhriati and E. Singh, “Descriptive Statistics in Business Research,” *Int. J. Adv. Res.*, vol. 3, no. 06, pp. 1409–1415, 2015.
- [48] C. shui Yu, T. ting Shao, and D. mo Li, “Distribution of standard deviation of an observable among superposed states,” *Ann. Phys. (N. Y.)*, vol. 373, pp. 43–51, 2016, doi:

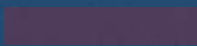
10.1016/j.aop.2016.06.017.

- [49] <sup>21</sup> D. N. Joanes and C. A. Gill, “Comparing measures of sample skewness and kurtosis,” *J. R. Stat. Soc. Ser. D Stat.*, vol. 47, no. 1, pp. 183–189, 1998, doi: 10.1111/1467-9884.00122.
- [50] <sup>37</sup> M. Xu and P. Shang, “Analysis of financial time series using multiscale entropy based on skewness and kurtosis,” *Phys. A: Stat. Mech. its Appl.*, vol. 490, pp. 1543–1550, 2018, doi: 10.1016/j.physa.2017.08.136.
- [51] <sup>38</sup> “Normality Testing - Skewness and Kurtosis | The GoodData Community.” <https://community.gooddata.com/metrics-and-maql-kb-articles-43/normality-testing-skewness-and-kurtosis-241> (accessed Jun. 15, 2021).
- [52] <sup>50</sup> M. J. Blanca, J. Arnau, D. López-Montiel, R. Bono, and R. Bendayan, “Skewness and kurtosis in real data samples,” *Methodology*, vol. 9, no. 2, pp. 78–84, 2013, doi: 10.1027/1614-2241/a000057.
- [53] <sup>68</sup> “Kurtosis - Definition, Excess Kurtosis, and Types of Kurtosis.” <https://corporatefinanceinstitute.com/resources/knowledge/other/kurtosis/> (accessed Jun. 15, 2021).
- [54] S. Narkhede, “Understanding Descriptive Statistics | by Sarang Narkhede | Towards Data Science,” 2018. <https://towardsdatascience.com/understanding-descriptive-statistics-c9c2b0641291> (accessed Jun. 12, 2021).
- [55] <sup>41</sup> J. W. Tukey, “Exploratory Data Analysis by John W. Tukey,” *Biometrics*, vol. 33, p. 768, 1977, [Online]. Available: <http://www.jstor.org/stable/2529486>.
- [56] <sup>15</sup> A. T. Jebb, S. Parrigon, and S. E. Woo, “Exploratory data analysis as a foundation of inductive research,” *Hum.*

*Resour. Manag. Rev.*, vol. 27, no. 2, pp. 265–276, 2017, doi: 10.1016/j.hrmr.2016.08.003.

- [57] C. Kneale and S. D. Brown, “Uncharted forest: A technique for exploratory data analysis,” *Talanta*, vol. 189, no. April, pp. 71–78, 2018, doi: 10.1016/j.talanta.2018.06.061.
- [58] S. Kelly, “What Is Python?,” in *Python, PyGame, and Raspberry Pi Game Development*, Berkeley, CA: Apress, 2019, pp. 5–9.
- [59] P. L. Gilabert *et al.*, “An efficient combination of digital predistortion and ofdm clipping for power amplifiers,” *Int. J. RF Microw. Comput. Eng.*, vol. 19, no. 5, pp. 583–591, 2009, doi: 10.1002/mmce.20381.
- [60] C. R. Harris *et al.*, “Array programming with NumPy,” *Nature*, vol. 585, no. 7875, pp. 357–362, 2020, doi: 10.1038/s41586-020-2649-2.
- [61] I. Stancin and A. Jovic, “An overview and comparison of free Python libraries for data mining and big data analysis,” *2019 42nd Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2019 - Proc.*, pp. 977–982, 2019, doi: 10.23919/MIPRO.2019.8757088.
- [62] R. Hubert, “Bureau of Aircraft Accidents Archives,” [Http://Www.Baaa-Acro.Com/](http://www.baaa-acro.com/), 2013. <http://www.baaa-acro.com/presentation/> (accessed May 28, 2021).
- [63] J. Frost, “Measures of Central Tendency: Mean, Median, and Mode - Statistics By Jim,” *Statistics by Jim*, 2018. <https://statisticsbyjim.com/basics/measures-central-tendency-mean-median-mode/> (accessed Jun. 11, 2021).





ISBN 978-979-587-991-6



# data analitik buku

## ORIGINALITY REPORT

16%

SIMILARITY INDEX

13%

INTERNET SOURCES

7%

PUBLICATIONS

9%

STUDENT PAPERS

## PRIMARY SOURCES

1	Submitted to Queen's University of Belfast Student Paper	2%
2	repository.its.ac.id Internet Source	1%
3	Submitted to University of Leicester Student Paper	1%
4	ojs.uho.ac.id Internet Source	<1%
5	Submitted to University of Glasgow Student Paper	<1%
6	Submitted to University of Sheffield Student Paper	<1%
7	www.researchgate.net Internet Source	<1%
8	Submitted to University of Wales, Lampeter Student Paper	<1%
9	repo.itera.ac.id Internet Source	<1%
10	eudl.eu Internet Source	<1%
11	Okyza Maherdy Prabowo, Suhono Harso Supangkat, Eueung Mulyana. "Anomaly Detection Techniques in Smart City: A Review from a Framework Perspective", 2021 International Conference on ICT for Smart Society (ICISS), 2021 Publication	<1%

12	<a href="https://ojs.pnb.ac.id">ojs.pnb.ac.id</a> Internet Source	<1 %
13	<a href="https://lindaaverett.homesandland.com">lindaaverett.homesandland.com</a> Internet Source	<1 %
14	<a href="https://edoc.ub.uni-muenchen.de">edoc.ub.uni-muenchen.de</a> Internet Source	<1 %
15	Submitted to Politeknik Negeri Bandung Student Paper	<1 %
16	Yuhao Wang, Yutian Pang, Oliver Chen, Hari N. Iyer, Parikshit Dutta, P.K. Menon, Yongming Liu. "Uncertainty quantification and reduction in aircraft trajectory prediction using Bayesian-Entropy information fusion", Reliability Engineering & System Safety, 2021 Publication	<1 %
17	Deris Stiawan, Ahmad Heryanto, Ali Berdadi, Dian Palupi Rini et al. "An Approach for Optimizing Ensemble Intrusion Detection Systems", IEEE Access, 2020 Publication	<1 %
18	Narender Kumar, Dharmender Kumar. "Machine Learning based Heart Disease Diagnosis using Non-Invasive Methods: A Review", Journal of Physics: Conference Series, 2021 Publication	<1 %
19	Submitted to Kaunas University of Technology Student Paper	<1 %
20	<a href="https://text-id.123dok.com">text-id.123dok.com</a> Internet Source	<1 %
21	<a href="https://www.biorxiv.org">www.biorxiv.org</a> Internet Source	<1 %
22	Submitted to Cranfield University Student Paper	<1 %

23	Submitted to National College of Ireland Student Paper	<1 %
24	docplayer.info Internet Source	<1 %
25	drepo.uniza.sk Internet Source	<1 %
26	Submitted to University of York Student Paper	<1 %
27	angkasasena.blogspot.com Internet Source	<1 %
28	ppid-dev.dephub.go.id Internet Source	<1 %
29	Submitted to Coleg Sir Gar Student Paper	<1 %
30	Submitted to Instituto Tecnológico de Aeronáutica Student Paper	<1 %
31	Submitted to Forum Komunikasi Perpustakaan Perguruan Tinggi Kristen Indonesia (FKPPTKI) Student Paper	<1 %
32	www.slideshare.net Internet Source	<1 %
33	Oliveira, J. M. B., P. F. Rodrigues, and H. M. Salgado. "Experimental performance assessment of a multicarrier digitised RoF system: Analysis and evaluation of pre-distortion techniques", 2014 16th International Conference on Transparent Optical Networks (ICTON), 2014. Publication	<1 %
34	Singh, A., A. Taneja, and G. Mangalaraj. "Creating online surveys: some wisdom from	<1 %

the trenches tutorial", IEEE Transactions on Professional Communication, 2009.

Publication

---

35 Anna Baron Garcia, Radu F. Babiceanu, Remzi Seker. "Artificial Intelligence and Machine Learning Approaches For Aviation Cybersecurity: An Overview", 2021 Integrated Communications Navigation and Surveillance Conference (ICNS), 2021

Publication

---

36 Submitted to Singapore Institute of Technology

Student Paper

---

37 Sou Nobukawa, Tomoaki Sekine, Masaru Chiba, Teruya Yamanishi, Haruhiko Nishimura. "Risk Analysis of Financial Time-Series Using Multi-Scale Entropy", 2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 2018

Publication

---

38 Submitted to University of Lincoln

Student Paper

---

39 [dblp.dagstuhl.de](http://dblp.dagstuhl.de)

Internet Source

---

40 [virgiliocervantes.co.uk](http://virgiliocervantes.co.uk)

Internet Source

---

41 [uvadoc.uva.es](http://uvadoc.uva.es)

Internet Source

---

42 [www.mii.lt](http://www.mii.lt)

Internet Source

---

43 [www.scribd.com](http://www.scribd.com)

Internet Source

---

44 [www.sunyrockland.edu](http://www.sunyrockland.edu)

Internet Source

<1 %

45

Submitted to Southern New Hampshire University - Continuing Education

Student Paper

<1 %

46

doku.pub

Internet Source

<1 %

47

ntrs.nasa.gov

Internet Source

<1 %

48

Submitted to Heriot-Watt University

Student Paper

<1 %

49

Madyaningarum Madyaningarum. "Evaluation of The Officer's Behaviour in Public Services of The Nuclear Minerals Technology", Jurnal Forum Nuklir, 2020

Publication

<1 %

50

Zhou, Qi, Junming Zhang, Jinglin Li, and Shangguang Wang. "Discovering Regional Taxicab Demand Based on Distribution Modeling from Trajectory Data", 2014 IEEE 17th International Conference on Computational Science and Engineering, 2014.

Publication

<1 %

51

www.ojs.stmikpringsewu.ac.id

Internet Source

<1 %

52

Abdhul Ahadh, Govind Vallabhasseri Binish, Rajagopalan Srinivasan. "Text Mining of Accident Reports using Semi-Supervised Keyword Extraction and Topic Modeling", Process Safety and Environmental Protection, 2021

Publication

<1 %

53

Janusz Kacprzyk, Anna Wilbik. "Chapter 47 A New Insight into the Linguistic Summarization

<1 %

of Time Series Via a Degree of Support:  
Elimination of Infrequent Patterns", Springer  
Science and Business Media LLC, 2008

Publication

54

Submitted to Nottingham Trent University

Student Paper

<1 %

55

repository.aust.edu.ng

Internet Source

<1 %

56

Ruihua Xu, Fan Luo, Gaoming Chen, Fenghua Zhou, Efraim Wozir Abdulahi. "Application of HFACS and grounded theory for identifying risk factors of air traffic controllers' unsafe acts", International Journal of Industrial Ergonomics, 2021

Publication

<1 %

57

id.scribd.com

Internet Source

<1 %

58

lib.unnes.ac.id

Internet Source

<1 %

59

dspace.mit.edu

Internet Source

<1 %

60

fisn.uni-plovdiv.bg

Internet Source

<1 %

61

hal.archives-ouvertes.fr

Internet Source

<1 %

62

Submitted to Higher Education Commission  
Pakistan

Student Paper

<1 %

63

Paloma Carollo Toscan, Alcindo Neckel, Laércio Stolfo Maculan, Cleiton Korcelski et al. "Use of geospatial tools to predict the risk of contamination by SARS-CoV-2 in urban cemeteries", Geoscience Frontiers, 2021

Publication

<1 %



64	Submitted to University of Stellenbosch, South Africa Student Paper	<1 %
65	<a href="http://eprints.uny.ac.id">eprints.uny.ac.id</a> Internet Source	<1 %
66	<a href="http://ujcontent.uj.ac.za">ujcontent.uj.ac.za</a> Internet Source	<1 %
67	Submitted to Blackpool and The Fylde College, Lancashire Student Paper	<1 %
68	Submitted to University of Salford Student Paper	<1 %
69	<a href="http://academic.oup.com">academic.oup.com</a> Internet Source	<1 %
70	<a href="http://e-repository.perpus.iainsalatiga.ac.id">e-repository.perpus.iainsalatiga.ac.id</a> Internet Source	<1 %
71	Marek Gagolewski. "Penalty-based aggregation of multidimensional data", Fuzzy Sets and Systems, 2017 Publication	<1 %
72	<a href="http://hdl.handle.net">hdl.handle.net</a> Internet Source	<1 %
73	<a href="http://repo.ikipgribali.ac.id">repo.ikipgribali.ac.id</a> Internet Source	<1 %
74	<a href="http://towardsdatascience.com">towardsdatascience.com</a> Internet Source	<1 %
75	<a href="http://www.tandfonline.com">www.tandfonline.com</a> Internet Source	<1 %
76	<a href="http://link.springer.com">link.springer.com</a> Internet Source	<1 %
77	<a href="http://onlinelibrary.wiley.com">onlinelibrary.wiley.com</a> Internet Source	<1 %

78	<a href="http://predatech.org">predatech.org</a> Internet Source	<1 %
79	<a href="http://psyjournals.ru">psyjournals.ru</a> Internet Source	<1 %
80	Aderibigbe Israel Adekitan. "Safeguards: A key process safety tool in jet fuel management from refinery to aircraft wings", Process Safety Progress, 2018 Publication	<1 %
81	Rosa María Arnaldo Valdés, V. Fernando Gómez Comendador, Luis Perez Sanz, Alvaro Rodriguez Sanz. "Prediction of aircraft safety incidents using Bayesian inference and hierarchical structures", Safety Science, 2018 Publication	<1 %
82	Submitted to Universitas Islam Indonesia Student Paper	<1 %
83	Submitted to University of Technology, Sydney Student Paper	<1 %
84	<a href="http://theses.gla.ac.uk">theses.gla.ac.uk</a> Internet Source	<1 %
85	<a href="http://umaralbasor46.wordpress.com">umaralbasor46.wordpress.com</a> Internet Source	<1 %
86	Submitted to Sriwijaya University Student Paper	<1 %
87	Submitted to Universitas Brawijaya Student Paper	<1 %
88	<a href="http://etheses.whiterose.ac.uk">etheses.whiterose.ac.uk</a> Internet Source	<1 %
89	<a href="http://issuu.com">issuu.com</a> Internet Source	<1 %

90	Internet Source	<1 %
91	Cattan, Mima. "Mental Health and Well-Being in Later Life", Mental Health and Well-Being in Later Life, 2009 Publication	<1 %
92	digitalcommons.lsu.edu Internet Source	<1 %
93	id.123dok.com Internet Source	<1 %
94	pt.slideshare.net Internet Source	<1 %
95	Submitted to Anglia Ruskin University Student Paper	<1 %
96	Submitted to National University of Singapore Student Paper	<1 %
97	Submitted to Universiti Teknologi MARA Student Paper	<1 %
98	arthinkle.com Internet Source	<1 %
99	core.ac.uk Internet Source	<1 %
100	dspace.uui.ac.id Internet Source	<1 %
101	etheses.uin-malang.ac.id Internet Source	<1 %
102	repository.ppns.ac.id Internet Source	<1 %
103	repository.unisba.ac.id Internet Source	<1 %
104	rmdopen.bmj.com Internet Source	<1 %

105	<a href="http://www.baliblogger.org">www.baliblogger.org</a> Internet Source	<1 %
106	<a href="http://www.buenastareas.com">www.buenastareas.com</a> Internet Source	<1 %
107	<a href="http://www.elettra.trieste.it">www.elettra.trieste.it</a> Internet Source	<1 %
108	<a href="http://www.idunn.no">www.idunn.no</a> Internet Source	<1 %
109	<a href="http://www.scirp.org">www.scirp.org</a> Internet Source	<1 %
110	<a href="http://www.siftdesk.org">www.siftdesk.org</a> Internet Source	<1 %
111	<a href="http://10her.dvrmyndns.com">10her.dvrmyndns.com</a> Internet Source	<1 %
112	Hyunseong Lee, Guoyi Li, Ashwin Rai, Aditi Chattopadhyay. "Real-time anomaly detection framework using a support vector regression for the safety monitoring of commercial aircraft", <i>Advanced Engineering Informatics</i> , 2020 Publication	<1 %
113	Putri Monica Sari, Moh. Bahrudin, Gustika Nurmalia. "STUDI KOMPARATIF ANALISIS EFISIENSI KINERJA PERBANKAN SYARIAH DI INDONESIA ANTARA METODE DATA ENVELOPMENT ANALYSIS (DEA) DAN STOCHASTIC FRONTIER ANALYSIS (SFA)", <i>FIDUSIA : JURNAL KEUANGAN DAN PERBANKAN</i> , 2020 Publication	<1 %
114	Submitted to Study Group Australia Student Paper	<1 %

115	Tebogo Makaba, Wesley Doorsamy, Babu Sena Paul. "Exploratory framework for analysing road traffic accident data with validation on Gauteng province data", Cogent Engineering, 2020 Publication	<1 %
116	Submitted to University of Surrey Student Paper	<1 %
117	Yusriza Mohamd Yusof, Afida Ayob, Mohamad Hanif Md Saad, Haryanti Mohd Affandi, Aini Hussain. "Need analysis for the development of a microcontroller instructional module programming literacy", Bulletin of Electrical Engineering and Informatics, 2020 Publication	<1 %
118	arxiv-export-lb.library.cornell.edu Internet Source	<1 %
119	citisee.amikompurwokerto.ac.id Internet Source	<1 %
120	deepai.org Internet Source	<1 %
121	digilib.uinsby.ac.id Internet Source	<1 %
122	diskominfo.sidoarjokab.go.id Internet Source	<1 %
123	eprints.uns.ac.id Internet Source	<1 %
124	es.scribd.com Internet Source	<1 %
125	jurnal.untan.ac.id Internet Source	<1 %
126	myleprac.blogspot.com Internet Source	<1 %

127	pt.scribd.com Internet Source	<1 %
128	repository.trisakti.ac.id Internet Source	<1 %
129	research.vumc.nl Internet Source	<1 %
130	utpedia.utp.edu.my Internet Source	<1 %
131	www.aanda.org Internet Source	<1 %
132	Prasanna Illankoon, Phillip Tretten, Uday Kumar. "A prospective study of maintenance deviations using HFACS-ME", International Journal of Industrial Ergonomics, 2019 Publication	<1 %
133	Tejas G. Puranik, Nicolas Rodriguez, Dimitri N. Mavris. "Towards nline prediction of safety-critical landing metrics in aviation using supervised machine learning", Transportation Research Part C: Emerging Technologies, 2020 Publication	<1 %
134	Vahid Zamanzadeh, Reza Ghaffari, Leila Valizadeh, Hossein Karimi-Moonaghi, Amy N.B. Johnston, Samaneh Alizadeh. "Challenges of objective structured clinical examination in undergraduate nursing curriculum: Experiences of faculties and students", Nurse Education Today, 2021 Publication	<1 %
135	Moustafa Mahmoud Yousry Zaghloul, Mai Mahmoud Yousry Zaghloul. "Influence of flame retardant magnesium hydroxide on the mechanical properties of high density polyethylene composites", Journal of Reinforced Plastics and Composites, 2017	<1 %

## Publication

---

Exclude quotes Off

Exclude bibliography Off

Exclude matches Off