

---

**KULLIYAH OF INFORMATION & COMMUNICATION TECHNOLOGY**

---

**INFO 4311 DATA WAREHOUSING**

**SEMESTER 1, 2024/2025**

**SECTION 01, GROUP A**

**GROUP PROJECT:  
DESIGN OF A STAR SCHEMA DATA WAREHOUSE  
Aviation Accident (2021 - 2023)**

**PREPARED BY:**

NO.	NAME	MATRIC NO
1	MUHAMMAD IKMAL HAKIMI BIN ROSLI	2210827
2	AHMAD IMAN BIN TURIN	2123349
3	MUHAMMAD MUSLIHUDDIN BIN MUSTAFFAR	2213263
4	ARMAN NURI ANUAR	2119279
5	ABDLGHANI AMER	1910115

**LECTURER**

ASSOC. PROF. DR. LILI MARZIANA ABDULLAH

**DUE**

21<sup>st</sup> NOVEMBER 2024

---

## TABLE OF CONTENTS

<b>1.0 BUSINESS PROCESS.....</b>	<b>1</b>
<b>2.0 KEY PERFORMANCE INDICATOR(s).....</b>	<b>2</b>
<b>3.0 DIMENSIONAL MODEL.....</b>	<b>2</b>
<b>4.0 BENEFITS.....</b>	<b>3</b>
<b>5.0 DETAILED DIMENSIONAL DESIGN.....</b>	<b>4</b>
5.1 DIMENSION TABLES.....	4
5.2 FACT TABLE.....	6
5.3 DESIGN OF A STAR SCHEMA.....	7

## 1.0 BUSINESS PROCESS

### Airline Safety Monitoring:

Tracking and analyzing airplane accidents to improve airline safety and operational efficiency. This process will leverage insights from the accident data to identify trends, reduce future accidents, and enhance resource allocation for safety protocols.

## 2.0 KEY PERFORMANCE INDICATOR(s)

1. **Total Number of Accidents:** Count of accidents over a time period.
2. **Fatality Rate:** Number of fatalities divided by total people aboard.
3. **Accidents by Aircraft Type:** Count of accidents grouped by aircraft type.
4. **Accidents by Operator:** Identify operators with the highest accident rates.
5. **Accidents by Location:** Identify high-risk areas.

## 3.0 DIMENSIONAL MODEL

Data Warehouse Bus Matrix

Business Process	Fact Table	Granularity	Facts	Date	Aircraft	Operator	Location	Accident Summary
Airline Safety Monitoring	PlaneAccidentFacts	One record per accident	AccidentCount, Fatalities, Aboard, Ground	X	X	X	X	X

### **Fact Table:**

- Fact Table Name : PlaneAccidentFacts
- Grain : One record per accident.
- Facts :
  - AccidentCount (Derived: always 1 per record).
  - Fatalities (Sum of fatalities in the accident).
  - Aboard (Total people aboard).
  - Ground (Number of ground casualties).

### **Dimension Tables:**

1. Date Dimension:
  - Helps analyze trends over time.
  - Columns: DateKey, FullDate, Year, Month, Day, Weekday.
2. Aircraft Dimension:
  - Details about the aircraft involved.
  - Columns: AircraftKey, AC\_Type, Registration, Construction\_no\_line\_no.
3. Operator Dimension:
  - Information about operators.
  - Columns: OperatorKey, OperatorName.
4. Location Dimension:
  - Details about the location of the accident.
  - Columns: LocationKey, LocationName.
5. Accident Summary Dimension:
  - Provides qualitative information about accidents.
  - Columns: SummaryKey, SummaryText.

## 4.0 BENEFITS

The dimensional model and KPIs would:

1. **Identify Trends:**

- Track high-risk areas, operators, and aircraft types.
- Analyze seasonal trends in accidents to improve scheduling and maintenance.

2. **Improve Safety Protocols:**

- Use fatality and ground casualty data to identify factors leading to severe accidents.
- Enhance operator safety training programs based on high-accident operators.

3. **Optimize Resources:**

- Focus resources on high-risk aircraft types, routes, and operators.

4. **Facilitate Decision-Making:**

- Provide insights for policymakers and regulatory authorities to enforce stricter safety measures.

## 5.0 DETAILED DIMENSIONAL DESIGN

### Plane Accident Dimensional Modelling

#### 5.1 DIMENSION TABLES

##### DimAircraft

Table Name	DimAircraft																		
Table Type	Dimension																		
Table Description	Provides details about the aircraft involved in events.																		
						Target					Source								
Column Name	Description	Unknown Member	Example Values	SCD Type	ETL Rule	Data Type	Size	Key	FK To	Null	Default Value	Source System	Source Table	Source Field Name	Source Data Type				
AircraftKey	Surrogate primary key for the aircraft	-1	1		Auto-generated sequential key	int		PK		N		Derived	Derived						
AC_Type	Type of the aircraft	-1	Boeing 737-524		Extract from source field	varchar				N		Plane Accidents Dataset	Plane Accidents Dataset	AC Type	Text				
Registration	Aircraft registration number	-1	PK-CLC		Extract from source field	varchar				N		Plane Accidents Dataset	Plane Accidents Dataset	Registration	Text				
Construction_no_line_no	Aircraft construction and line number	-1	27323/2616		Extract from source field	varchar				N		Plane Accidents Dataset	Plane Accidents Dataset	Construction no / line no	Text				

##### DimDate

Table Name :	DimDate																		
Table Type :	Dimension																		
Table Description	Standard Date Dimension																		
						Target						Source							
Column Name	Description	Unknown Member	Example Values	SCD Type	ETL Rule	Data Type	Size	Key	FK To	Null	Default Value	Source System	Source Table	Source Field Name	Source Data Type				
DateKey	Surrogate primary key YYYYMMDD	-1	20210109		Convert Date in datetime to int of format YYYYMMDD	int		PK		N		Derived	Derived	Date	DateTime				
FullDate	Full date (business key)	-1	9/1/2021		Use Date as-is	datetime				N		Derived	Derived	Date	DateTime				
Year	Year of the date	0	2021		Use function Year to extract Year	int				N		Derived	Derived	Date	DateTime				
Month	Month of the date	0	1		Use function Month to extract Month	int				N		Derived	Derived	Date	DateTime				
Day	Day part of the date	0	9		Use function Day to extract Day	int				N		Derived	Derived	Date	DateTime				

## DimLocation

<b>Table Name :</b>	DimLocation													
<b>Table Type :</b>	Dimension													
<b>Table Descriptio Details of Location</b>														
Column Name	Description	Unknown Member	Example Values	ETL Rule	Data Type	Size	Target				Source			
							Key	FK To	Null	Default Value	Source System	Source Table	Source Field Name	Source Data Type
LocationKey	Surrogate primary key for location	-1	1	Auto-generated sequential key	int		PK		N		Derived			
LocationName	Name of the location	-1	Near Jakarta, Indonesia	Extract from source field	nvarchar	50			N		Plane Accidents Dataset	2021 Plane Accident, 2022 Plane Accident, 2023 Plane Accident	Location	nvarchar

## DimOperator

Table Name :	DimOperator													
Table Type :	Dimension													
Table Description	Details of operator													
Column Name	Description	Unknown Member	Example Values	Target							Source			
				ETL Rule	Data Type	Size	Key	FK To	Null	Default Value	Source System	Source Table	Source Field Name	Source Data Type
OperatorKey	Surrogate primary key for operator	-1	1	Auto-generated sequential key	int		PK		N		Derived			
OperatorName	Name of the operator	-1	Sriwijaya Air	Extract from source field	nvarchar	50			N		Plane Accidents Dataset	2021 Plane Accident, 2022 Plane Accident, 2023 Plane Accident	Operator	nvarchar

### DimSummary

Table Name :	DimSummary													
Table Type :	Dimension													
Table Description	Accident Summary													
				Target							Source			
Column Name	Description	Unknown Member	Example Values	ETL Rule	Data Type	Size	Key	FK To	Null	Default Value	Source System	Source Table	Source Field Name	Source Data Type
SummaryKey	Surrogate primary key for summary	-1	1	Auto-generated sequential key	int		PK		N		Derived			
SummaryText	Text of the summary	-1	One of the engines on the aircraft failed 10 minutes after takeof. When the plane turned back, the second engine failed.	Extract from source field	nvarchar	700			N		Plane Accidents Dataset	2021 Plane Accident, 2022 Plane Accident, 2023 Plane Accident	Summary	nvarchar

## 5.2 FACT TABLE

## PlaneAccidentFacts

Table Name :	PlaneAccidentFacts																			
Table Type :	Fact																			
Table Description :	Each record show the details of an accident																			
Column Name	Description	Unknown Member	Example Values	SCD Type	ETL Rule	Target Data Type	Size	Key	FK To	Null	Default Value	Source Source System	Source Table	Source Field Name	Source Data Type					
AccidentKey	Surrogate primary key for each accident	-1	1		Auto-generated	int		PK		N		Derived								
DateKey	FK to DimDate	-1	20210109		Surrogate key pipeline lookup for dimension key lookup using	datetime		FK	DimDate.DateKey	N		DW	DimDate	DateKey	datetime					
AircraftKey	FK to DimAircraft	-1	1		Surrogate key pipeline lookup for dimension key lookup using	int		FK	DimAircraft.AircraftKey	N		DW	DimAircraft	AircraftKey	int					
LocationKey	FK to DimLocation	-1	1		Surrogate key pipeline lookup for dimension key lookup using	int		FK	DimLocation.LocationKey	N		DW	DimLocation	LocationKey	int					
OperatorKey	FK to DimOperator	-1	1		Surrogate key pipeline lookup for dimension key lookup using	int		FK	DimOperator.OperatorKey	N		DW	DimOperator	OperatorKey	int					
SummaryKey	FK to DimSummary	-1	1		Surrogate key pipeline lookup for dimension key lookup using	int		FK	DimSummary.SummaryKey	N		DW	DimSummary	SummaryKey	int					
AccidentCount	Always 1 (Represent an accident)	-1	1		Hardcoded	int				N	1	Derived								
Fatalities	Total fatalities in an accident	-1	62		Extract from source field as it is	int				Y		Plane Accidents Dataset	2021 Plane Accident, 2022 Plane Accident, 2023 Plane	Fatalities	int					
Aboard	Total aboard in an accident	-1	62		Extract from source field as it is	int				Y		Plane Accidents Dataset	2022 Plane Accident, 2022 Plane Accident, 2023 Plane	Aboard	int					
Ground	Total ground in accident	-1	0		Extract from source field as it is	int				Y		Plane Accidents Dataset	2023 Plane Accident, 2022 Plane Accident, 2023 Plane Accident	Ground	int					



## 5.3 DESIGN OF A STAR SCHEMA

