Project Proposal

Group 33
Vidit Kothari vk3080
Suraj Sureshkumar ss7495
Shlok Gupta sg6935

Datasets:

1. Openflights dataset

Link - https://openflights.org/data.html
This dataset has multiple tables that we will be using.

a. Airports:
   This is a table which has all the airports with their IATA and ICAO codes. The table also contains latitudes, longitudes, altitude and the timezone of the airport
b. Airlines:
   The table has information about all airlines, pulled from the wikipedia list of airlines. It has the airline name, the alias, the IATA code, ICAO code, country where the airline is based out of, and the callsign
c. Countries:
   This is a table with the country names and the ISO codes
d. Routes:
   The table contains all route information with data about the source, destination, airline, codeshare, stops and equipment used.
e. Aircrafts:
   This is a list of all equipment with their IATA and ICAO codes.

2. The airline delays dataset

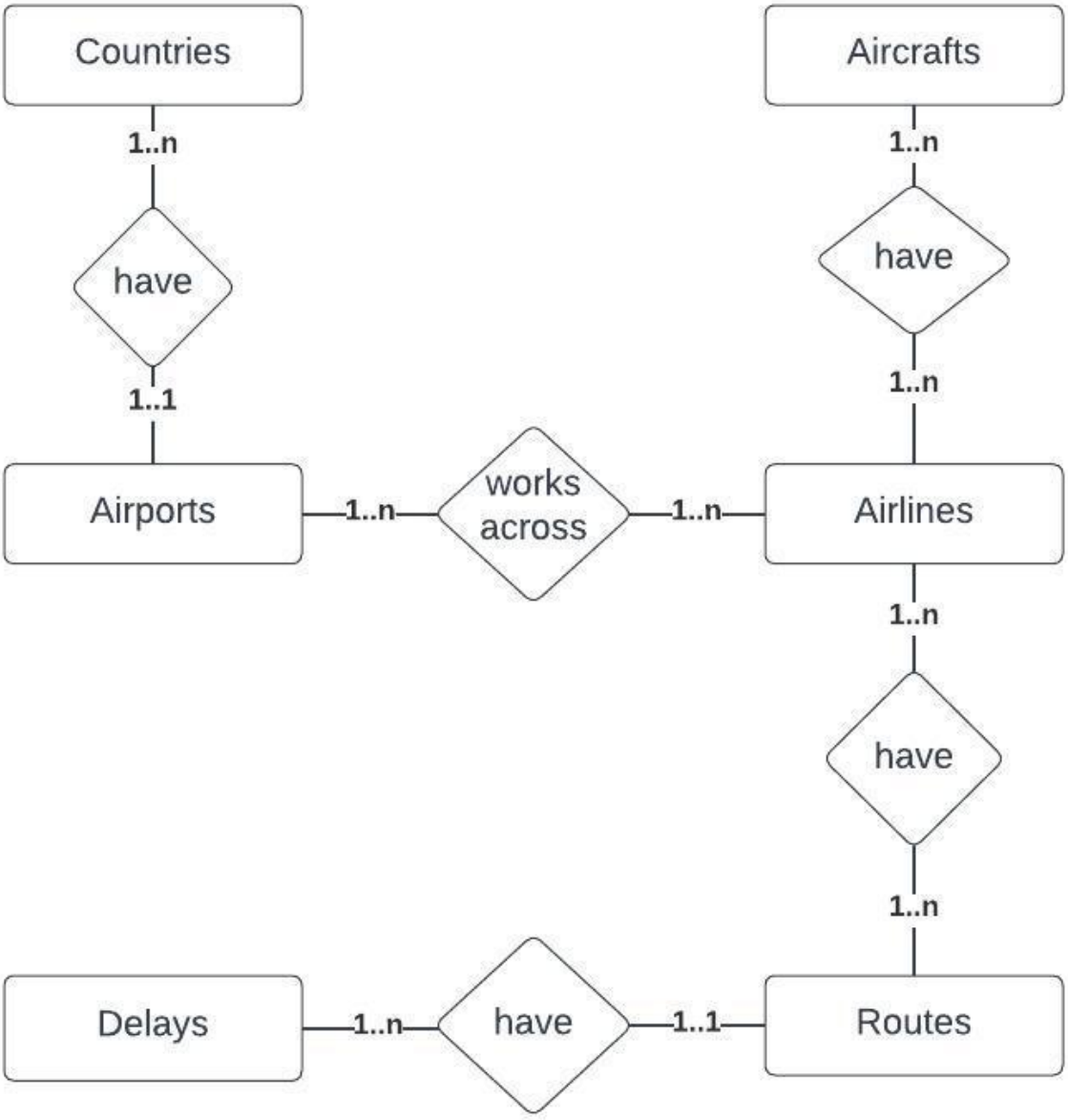
Link - https://www.kaggle.com/datasets/yuanyuwendymu/airline-delay-and-cancellation-data-2009-2018
This dataset has tables of delays in the US from 2009 to 2018. However, since the openflights dataset has data only till 2013, we will be using the data only from 2009 to 2013.
Each table contains information about the source and destination of the route, the airline operating it, the date of departure, the flight number, planned and

actual departure times, the total delay, the taxi time, the flight time, scheduled and actual arrival time, cancellation reason if the flight was canceled,and the reasons for delay.

3. The crashes dataset

Link - http://www.planecrashinfo.com/database.htm
We are also considering using this dataset as we think we may be able to derive meaningful associations between the crashes and delays.
This dataset contains information about crashes which include the date of accident,time,airline,flight number,aircraft type, ICAO code of the aircraft,total passengers and crew aboard, total fatalities,and a brief description of the accident.

The ER Diagram



Creating Tables

```
create table flights.airlines
(
    airline_id   integer not null
        constraint airlines_pk
            primary key,
    airline_name varchar,
    alias        varchar,
    "IATA_code"  varchar not null,
    "ICAO_code"  varchar not null,
    callsign     varchar,
    country      varchar,
    active       varchar
);



create table flights."Airport"
(
    airport_id   integer not null
        constraint airport
            primary key,
    airline_name varchar,
    city         varchar,
    country      varchar,
    "IATA_code"  varchar not null,
    "ICAO_code"  varchar not null,
    latitude     integer,
    longitude    integer,
    altitude     integer,
    timezone     integer,
    "DST_zone"   varchar
);


create table flights.countries
(
    country_name varchar,
    iso_code     varchar not null,
    country_id   integer not null
        constraint countries_pk
            primary key
);
```

```
create table flights.planes
(
    aircraft_name varchar,
    "IATA_code"   varchar,
    "ICAO_code"   varchar,
    plane_id      integer not null
        constraint planes_pk
            primary key
);


create table flights.routes
(
    airline        varchar,
    id             integer
        constraint routes_airlines_airline_id_fk
            references flights.airlines,
    source         varchar,
    source_id      integer,
    destination    varchar,
    destination_id integer,
    codeshare      integer,
    stops          integer,
    equipment      integer
);


create table flights.airlines_aircrafts
(
    airlines_id integer
        constraint airlines_aircrafts_airlines_airline_id_fk
            references flights.airlines,
    plane_id    integer
        constraint airlines_aircrafts_planes_plane_id_fk
            references flights.planes
);


create table flights.countries_airport
(
    airport_id integer
        constraint countries_airport_fk
```

```
        references flights."Airport",
    country_id integer
        constraint country_id_fk
            references flights.countries
);
```

Data loading

We used the import tool in datagrip to load the data.
Here are the snippets of the data :

1. Airports

	airport_id	airport_name	city	country	IATA_code	ICAO_code	latitude	longitude	a
1	1	Goroka Airport	Goroka	Papua New Guinea	GKA	AYGA	-6	145	
2	2	Madang Airport	Madang	Papua New Guinea	MAG	AYMD	-5	145	
3	3	Mount Hagen Kagamuga Airport	Mount Hagen	Papua New Guinea	HGU	AYMH	-5	144	
4	4	Nadzab Airport	Nadzab	Papua New Guinea	LAE	AYNZ	-6	146	
5	5	Port Moresby Jacksons International Airport	Port Moresby	Papua New Guinea	POM	AYPY	-9	147	
6	6	Wewak International Airport	Wewak	Papua New Guinea	WWK	AYWK	-3	143	
7	7	Narsarsuaq Airport	Narssarssuaq	Greenland	UAK	BGBW	61	-45	
8	8	Godthaab / Nuuk Airport	Godthaab	Greenland	GOH	BGGH	64	-51	
9	9	Kangerlussuaq Airport	Sondrestrom	Greenland	SFJ	BGSF	67	-50	
10	10	Thule Air Base	Thule	Greenland	THU	BGTL	76	-68	
11	11	Akureyri Airport	Akureyri	Iceland	AEY	BIAR	65	-18	
12	12	Egilsstaðir Airport	Egilsstadir	Iceland	EGS	BIEG	65	-14	
13	13	Hornafjörður Airport	Hofn	Iceland	HFN	BIHN	64	-15	
14	14	Húsavik Airport	Husavik	Iceland	HZK	BIHU	65	-17	
15	15	Ísafjörður Airport	Isafjordur	Iceland	IFJ	BIIS	66	-23	
16	16	Keflavik International Airport	Keflavik	Iceland	KEF	BIKF	63	-22	

2. Aircrafts

	aircraft_name	IATA_code	ICAO_code	plane_id
1	Aerospatiale (Nord) 262	ND2	N262	0
2	Aerospatiale (Sud Aviation) Se.210 Caravelle	CRV	S210	1
3	Aerospatiale SN.601 Corvette	NDC	S601	2
4	Aerospatiale/Alenia ATR 42-300	AT4	AT43	3
5	Aerospatiale/Alenia ATR 42-500	AT5	AT45	4
6	Aerospatiale/Alenia ATR 42-600	ATR	AT46	5
7	Aerospatiale/Alenia ATR 72	AT7	AT72	6
8	Airbus A300	AB3	A30B	7
9	Airbus A300-600	AB6	A306	8
10	Airbus A300-600ST Super Transporter / Beluga	ABB	A3ST	9
11	Airbus A310	310	A310	10
12	Airbus A318	318	A318	11
13	Airbus A319	319	A319	12
14	Airbus A319neo	31N	A19N	13
15	Airbus A320	320	A320	14
16	Airbus A320neo	32N	A20N	15
17	Airbus A321	321	A321	16

3. Countries

	country_name	iso_code	country_id
1	Bonaire, Saint Eustatius and Saba	BQ	0
2	Aruba	AW	1
3	Antigua and Barbuda	AG	2
4	United Arab Emirates	AE	3
5	Afghanistan	AF	4
6	Algeria	DZ	5
7	Azerbaijan	AZ	6
8	Albania	AL	7
9	Armenia	AM	8
10	Angola	AO	9
11	American Samoa	AS	10
12	Argentina	AR	11
13	Australia	AU	12
14	Ashmore and Cartier Islands	\N	13
15	Austria	AT	14
16	Anguilla	AI	15
17	Antarctica	AQ	16

4. Airlines

	airline_id	airline_name	alias	IATA_code	ICAO_code	callsign	country
1	-1	Unknown	\N	-	N/A	\N	\N
2	1	Private flight	\N	-	N/A		
3	2	135 Airways	\N		GNL	GENERAL	United States
4	3	1Time Airline	\N	1T	RNX	NEXTIME	South Africa
5	4	2 Sqn No 1 Elementary Flying Training School	\N		WYT		United Kingdom
6	5	213 Flight Unit	\N		TFU		Russia
7	6	223 Flight Unit State Airline	\N		CHD	CHKALOVSK-AVIA	Russia
8	7	224th Flight Unit	\N		TTF	CARGO UNIT	Russia
9	8	247 Jet Ltd	\N		TWF	CLOUD RUNNER	United Kingdom
10	9	3D Aviation	\N		SEC	SECUREX	United States
11	10	40-Mile Air	\N	Q5	MLA	MILE-AIR	United States
12	11	4D Air	\N		QRT	QUARTET	Thailand
13	12	611897 Alberta Limited	\N		THD	DONUT	Canada
14	13	Ansett Australia	\N	AN	AAA	ANSETT	Australia
15	14	Abacus International	\N	1B			Singapore
16	15	Abelag Aviation	\N	W9	AAB	ABG	Belgium

5. Routes

	airline	id	source	source_id	destination_id	codeshare	stops	equipment	destination
1	2B	410	AER	2965	2990		0	CR2	KZN
2	2B	410	ASF	2966	2990		0	CR2	KZN
3	2B	410	ASF	2966	2962		0	CR2	MRV
4	2B	410	CEK	2968	2990		0	CR2	KZN
5	2B	410	CEK	2968	4078		0	CR2	OVB
6	2B	410	DME	4029	2990		0	CR2	KZN
7	2B	410	DME	4029	6969		0	CR2	NBC
8	2B	410	DME	4029	<null>		0	CR2	TGK
9	2B	410	DME	4029	6160		0	CR2	UUA
10	2B	410	EGO	6156	2952		0	CR2	KGD
11	2B	410	EGO	6156	2990		0	CR2	KZN
12	2B	410	GYD	2922	6969		0	CR2	NBC
13	2B	410	KGD	2952	6156		0	CR2	EGO
14	2B	410	KZN	2990	2965		0	CR2	AER
15	2B	410	KZN	2990	2966		0	CR2	ASF
16	2B	410	KZN	2990	2968		0	CR2	CEK
17	2B	410	KZN	2990	4029		0	CR2	DME