Detailed Report Document:

Aviation Data Analysis

Executive Summary:

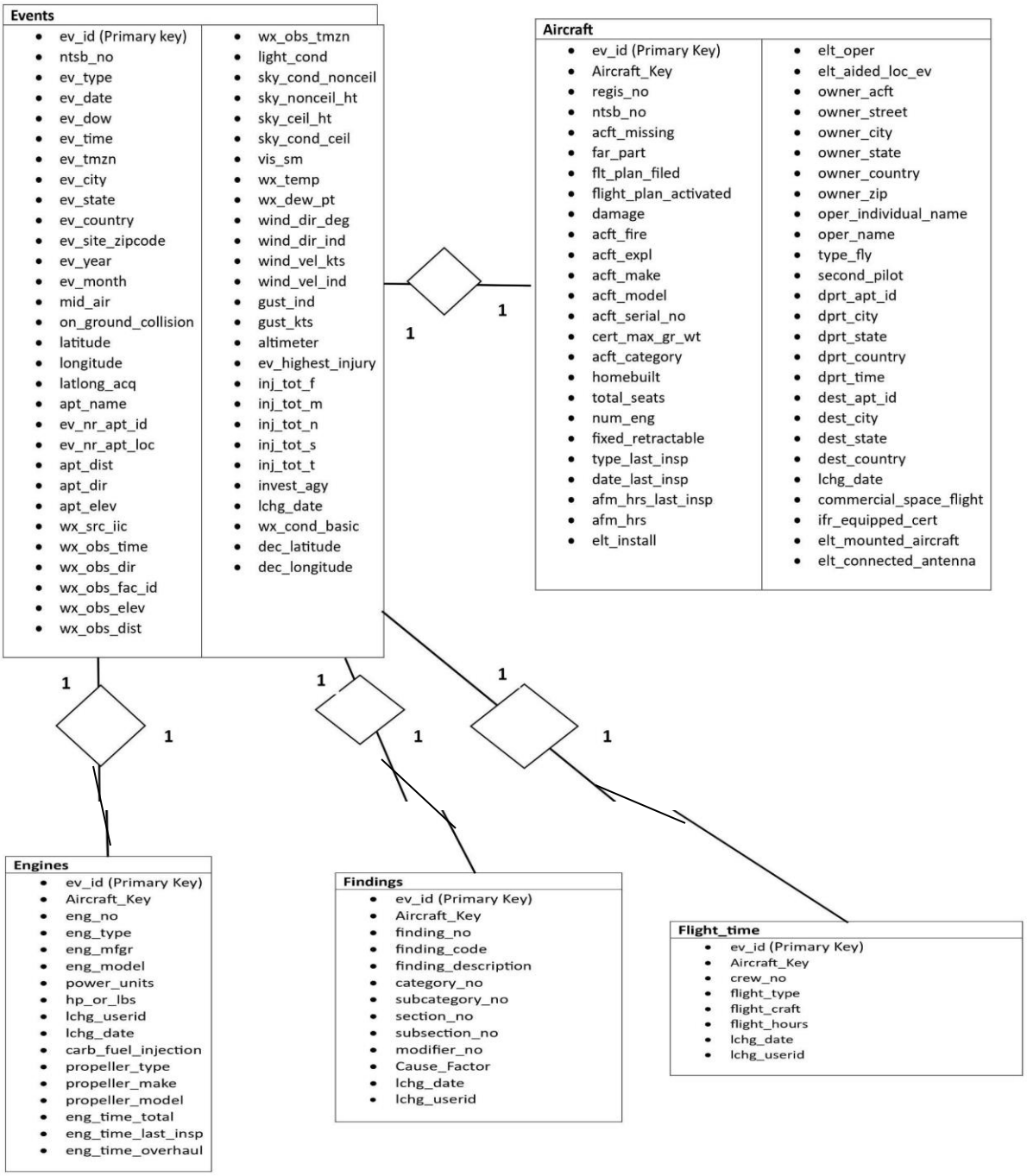
Objectives:

- Over the past fourteen years, aviation accident rates have shown limited improvement.
- This has resulted in a continued occurrence of injuries within the industry.
- Despite ongoing safety efforts, accidents have not significantly decreased.
- This trend emphasizes the need for continued vigilance and safety enhancement in aviation.

Key Problems:

Over the past fourteen years, the aviation industry has witnessed a concerning trend where the rate of aviation accidents and resulting injuries has remained relatively stable, showing limited signs of improvement.

Entity Relationship Diagram (ERD):


Events
ev_id (Primary key)
ntsb_no
ev_type
ev_date
ev_dow
ev_time
ev_tmzn
ev_city
ev_state
ev_country
ev_site_zipcode
ev_year
ev_month
mid_air
on_ground_collision
latitude
longitude
latlong_acq
apt_name
ev_nr_apt_id
ev_nr_apt_loc
apt_dist
apt_dir
apt_elev
wx_src_iic
wx_obs_time
wx_obs_dir
wx_obs_fac_id
wx_obs_elev
wx_obs_dist
wx_obs_tmzn
light_cond
sky_cond_nonceil
sky_nonceil_ht
sky_ceil_ht
sky_cond_ceil
vis_sm
wx_temp
wx_dew_pt
wind_dir_deg
wind_dir_ind
wind_vel_kts
wind_vel_ind
gust_ind
gust_kts
altimeter
ev_highest_injury
inj_tot_f
inj_tot_m
inj_tot_n
inj_tot_s
inj_tot_t
invest_agy
lchg_date
wx_cond_basic
dec_latitude
dec_longitude
Aircraft
ev_id (Primary Key)
Aircraft_Key
regis_no
ntsb_no
acft_missing
far_part
flt_plan_filed
flight_plan_activated
damage
acft_fire
acft_expl
acft_make
acft_model
acft_serial_no
cert_max_gr_wt
acft_category
homebuilt
total_seats
num_eng
fixed_retractable
type_last_insp
date_last_insp
afm_hrs_last_insp
afm_hrs
elt_install
elt_oper
elt_aided_loc_ev
owner_acft
owner_street
owner_city
owner_state
owner_country
owner_zip
oper_individual_name
oper_name
type_fly
second_pilot
dprt_apt_id
dprt_city
dprt_state
dprt_country
dprt_time
dest_apt_id
dest_city
dest_state
dest_country
lchg_date
commercial_space_flight
ifr_equipped_cert
elt_mounted_aircraft
elt_connected_antenna
1
1
1
1
1
1
1
1
1
Engines
ev_id (Primary Key)
Aircraft_Key
eng_no
eng_type
eng_mfgr
eng_model
power_units
hp_or_lbs
lchg_userid
lchg_date
carb_fuel_injection
propeller_type
propeller_make
propeller_model
eng_time_total
eng_time_last_insp
eng_time_overhaul
Findings
ev_id (Primary Key)
Aircraft_Key
finding_no
finding_code
finding_description
category_no
subcategory_no
section_no
subsection_no
modifier_no
Cause_Factor
lchg_date
lchg_userid
Flight_time
ev_id (Primary Key)
Aircraft_Key
crew_no
flight_type
flight_craft
flight_hours
lchg_date
lchg_userid

Objectives and Approach:

Define objectives:

Analyze aviation data to check the trend in the enhancement of safety and give recommendations to improve safety on the aviation system at different points.

Stakeholders:

- Regulatory bodies like the Federal Aviation Administration (FAA) in the United States or the European Union Aviation Safety Agency (EASA) may be interested in findings like this.
- Airport Authorities play a critical role in aviation because they monitor air traffic.

Data Analysis Process / Approach:

- Understanding Key Problems and Define Objectives
- Data Preparation and Cleaning
- Data Processing and Analysis
- Visualization, generating insights and Recommendations

Data Exploration:

- Analyzed trends over different geographical locations.
- Analyzed month-on-month and year-on-year trends.
- Analyzed trends over different kinds of aircraft.

Questions, SQL Queries, Visualization, Results and Insights:

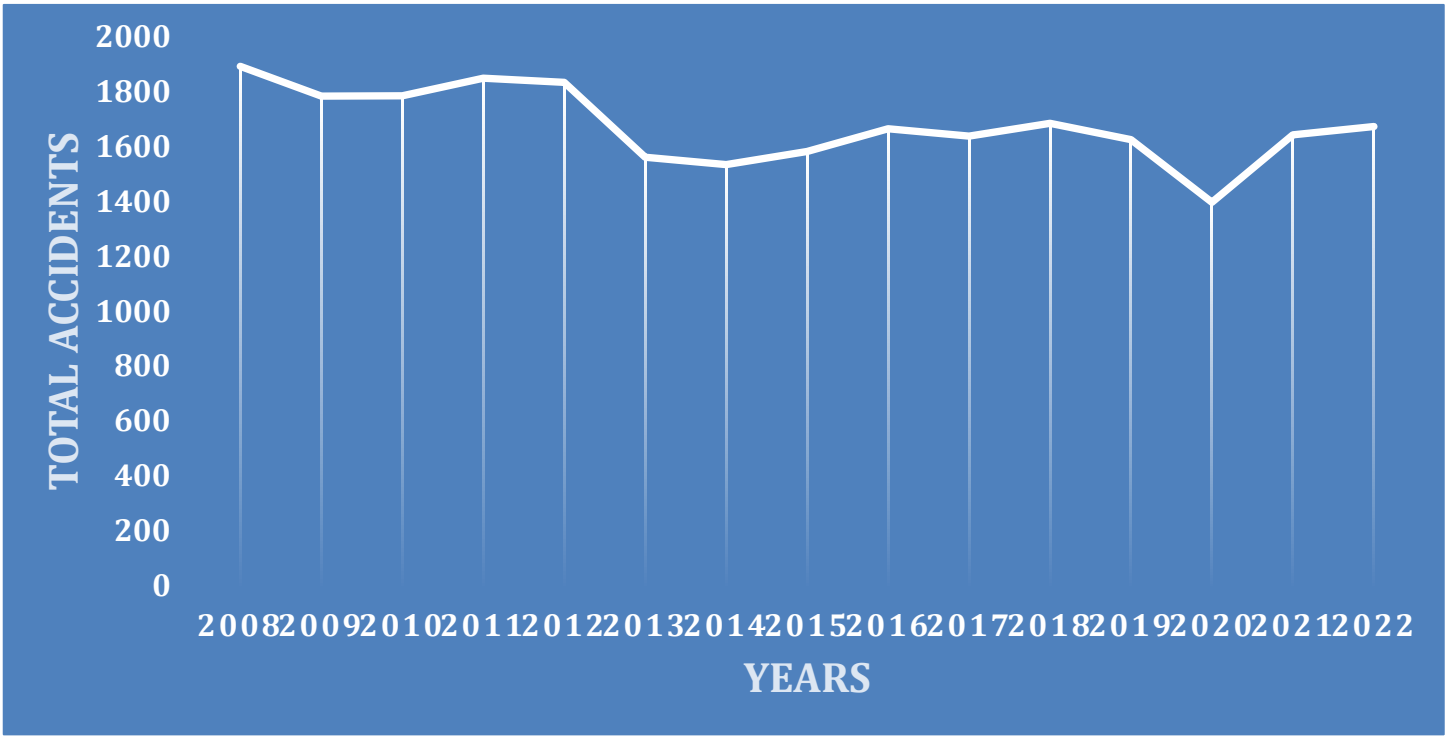
Q No 1: What is the number of total accidents in each year?

```
select count(*) as total_events, ev_year from events
group by ev_year order by ev_year ;
```

150 %

Results | Messages

	total_events	ev_year
1	1893	2008
2	1784	2009
3	1786	2010
4	1850	2011
5	1835	2012
6	1561	2013
7	1535	2014
8	1582	2015
9	1665	2016
10	1638	2017
11	1685	2018
12	1625	2019
13	1397	2020
14	1643	2021
15	1673	2022
16	886	2023



Aviation accidents per year have just a tiny dip over the last few years with some how the same rate over the

decade.

Although there was a nominal around the year 2020 due to low air space activities due to COVID-19.

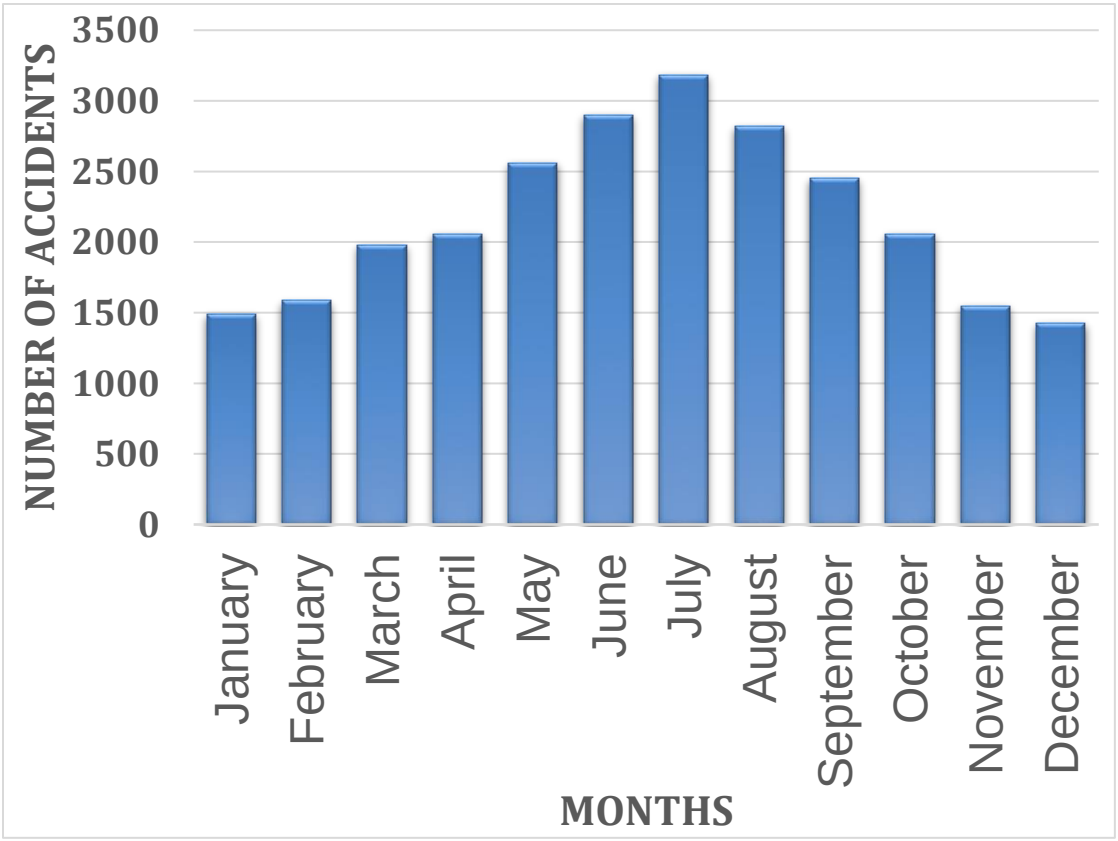
Q NO 2: In which months were the total accidents occurred?

```
select count(*) as total_events , month(event_date)
from events GROUP BY month(event_date) ORDER BY month(event_date);
```

50 %

Results | Messages

	total_events	(No column name)
1	1487	1
2	1584	2
3	1979	3
4	2054	4
5	2557	5
6	2900	6
7	3183	7
8	2819	8
9	2450	9
10	2054	10
11	1545	11
12	1426	12



The number of events increases during the mid-year months, as July seems to be the month with the highest count of accidents.

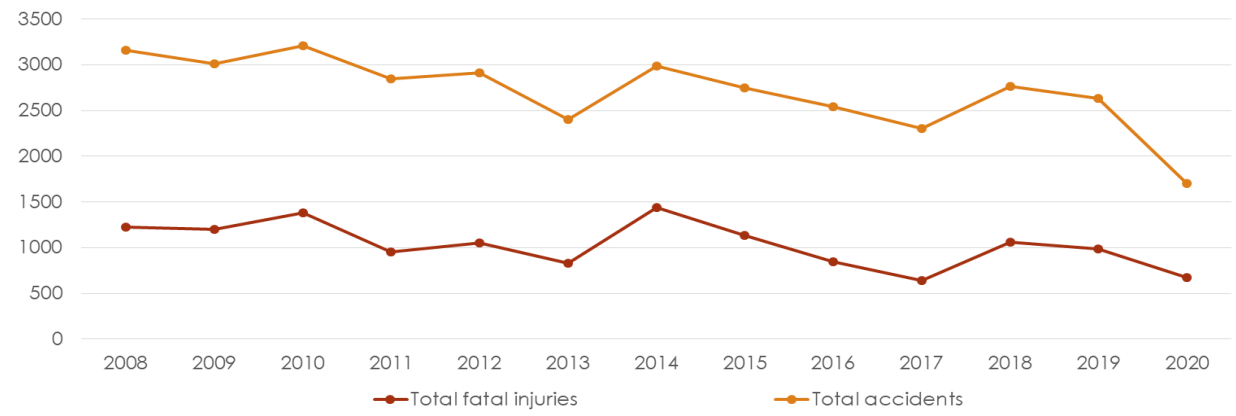
Q NO 3: What is the trend of total injuries with accidents per year?

```
select count(*) as total_events, sum(inj_tot_t) as total_injuries, ev_year from events
group by ev_year order by ev_year ;
```

150 %

Results | Messages

	total_events	total_injuries	ev_year
1	1893	13026	2008
2	1784	12419	2009
3	1786	13418	2010
4	1850	16260	2011
5	1835	13629	2012
6	1561	11028	2013
7	1535	11960	2014
8	1582	11653	2015
9	1665	11747	2016
10	1638	13396	2017
11	1685	13290	2018
12	1625	11955	2019
13	1397	6333	2020
14	1643	9390	2021
15	1673	12672	2022
16	886	10905	2023



The total number of fatal injuries also seems to be don't have much difference from the trend of accidents. But the graph for fatal injuries shows much improvement after the year 2018.

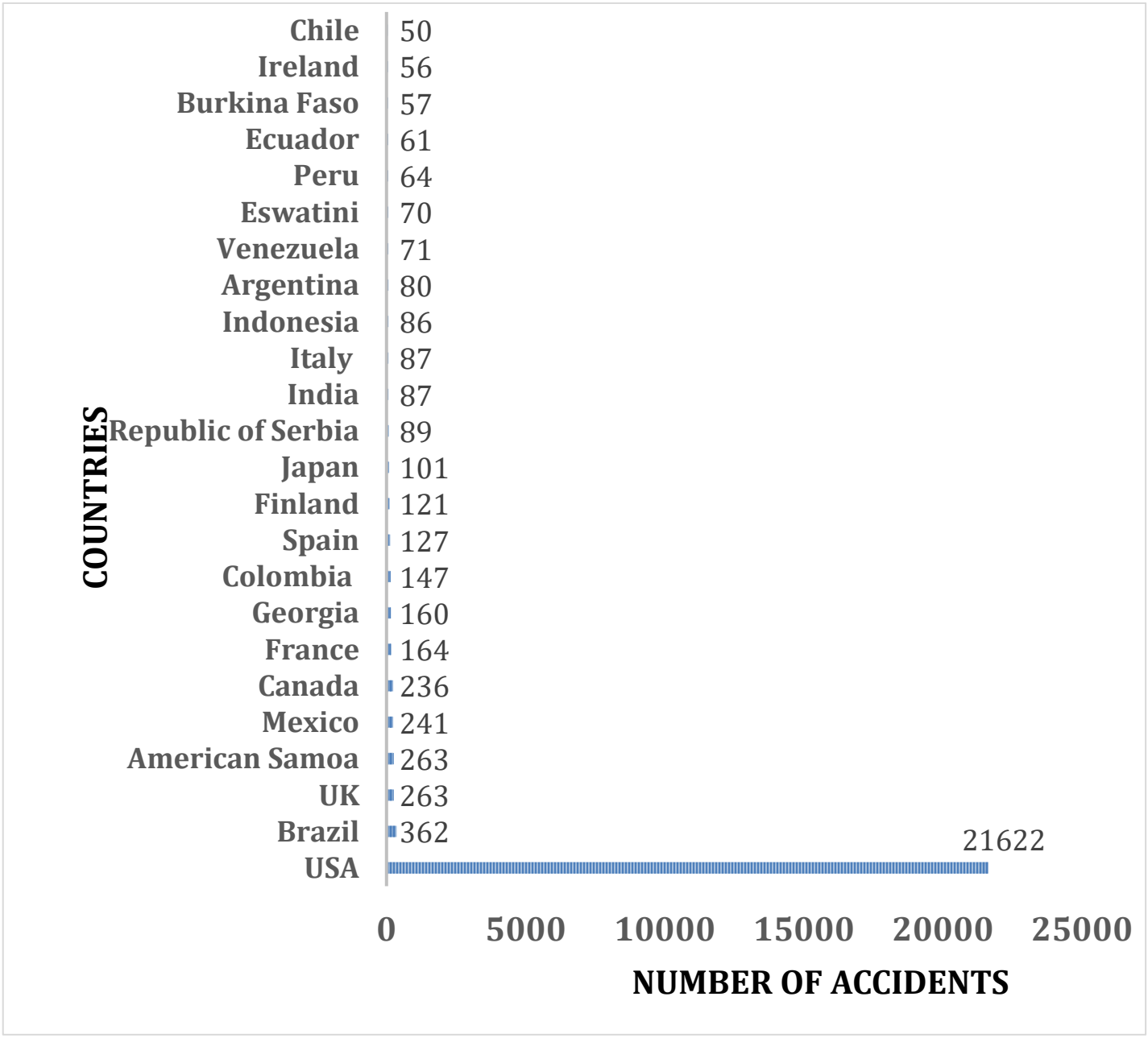
Q No 4 In which country were the most accidents reported?

```
--count(events.ev_id) as event_count, events.ev_year
--from events
--join aircraft on events.ev_id = aircraft.ev_id
--group by ev_year order by ev_year;

select count(*) as total_events, ev_country
from events group by ev_country order by total_events desc;

--select count(*) as aircraft_count, aircraft.acft_category
--from events join aircraft on events.ev_id = aircraft.ev_id
```

150 %

Results | Messages

	total_events	ev_country
1	21622	USA
2	362	BR
3	263	UK
4	263	AS
5	241	MX



Of the total **26038 accidents**, alone the **USA** is the country in which air space is reported with the highest number of events with no one even near it.

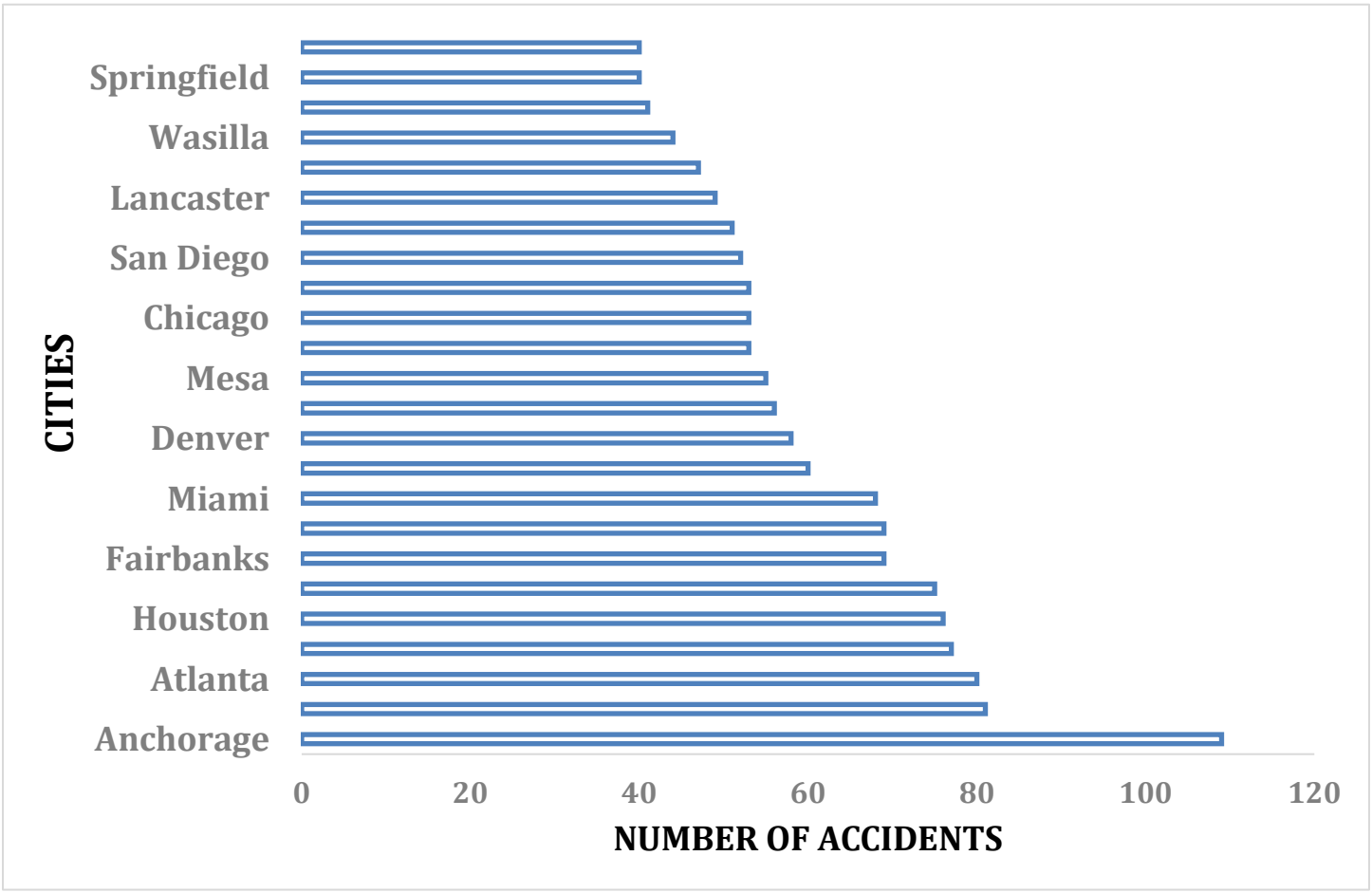
Q No 5: Which city observed the highest number of accidents?

```
select count(*) as total_events, ev_city from events
group by ev_city order by total_events desc;
```

150 %

Results | Messages

	total_events	ev_city
1	109	Anchorage
2	81	Phoenix
3	80	Atlanta
4	77	Palmer
5	76	Houston
6	75	Talkeetna
7	69	Fairbanks
8	69	Las Vegas
9	68	Miami
10	60	Reno
11	58	Denver
12	56	San Antonio
13	55	Mesa
14	53	London
15	53	Chicago
16	53	Albuquerque
17	52	San Diego



In total air spaces of **9154 cities**, aviation accidents were recorded.

The Alaskan city **Anchorage** with the most number of events reported.

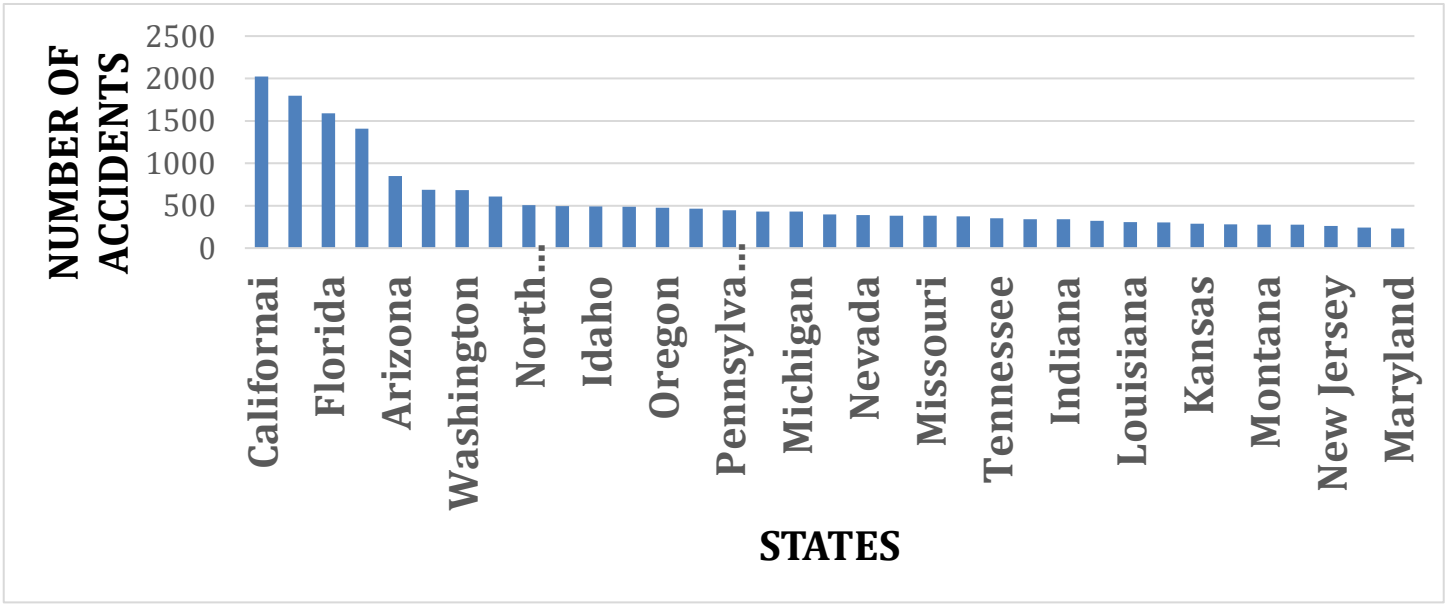
Q No 6: Which states of the USA have observed the most number of accidents?

```
select count(*) as total_events, ev_state from events where ev_country = 'USA'
group by ev_state order by total_events desc;
```

150 %

Results | Messages

	total_events	ev_state
1	2024	CA
2	1796	TX
3	1590	FL
4	1410	AK
5	848	AZ
6	686	CO
7	682	WA
8	609	GA
9	505	NC
10	494	IL
11	493	ID
12	488	NY
13	475	OR
14	464	OH
15	446	PA
16	431	UT
17	430	MI



California, **Texas**, **Florida,** and **Alaska** has the highest number of accidents with events in four figures were reported.

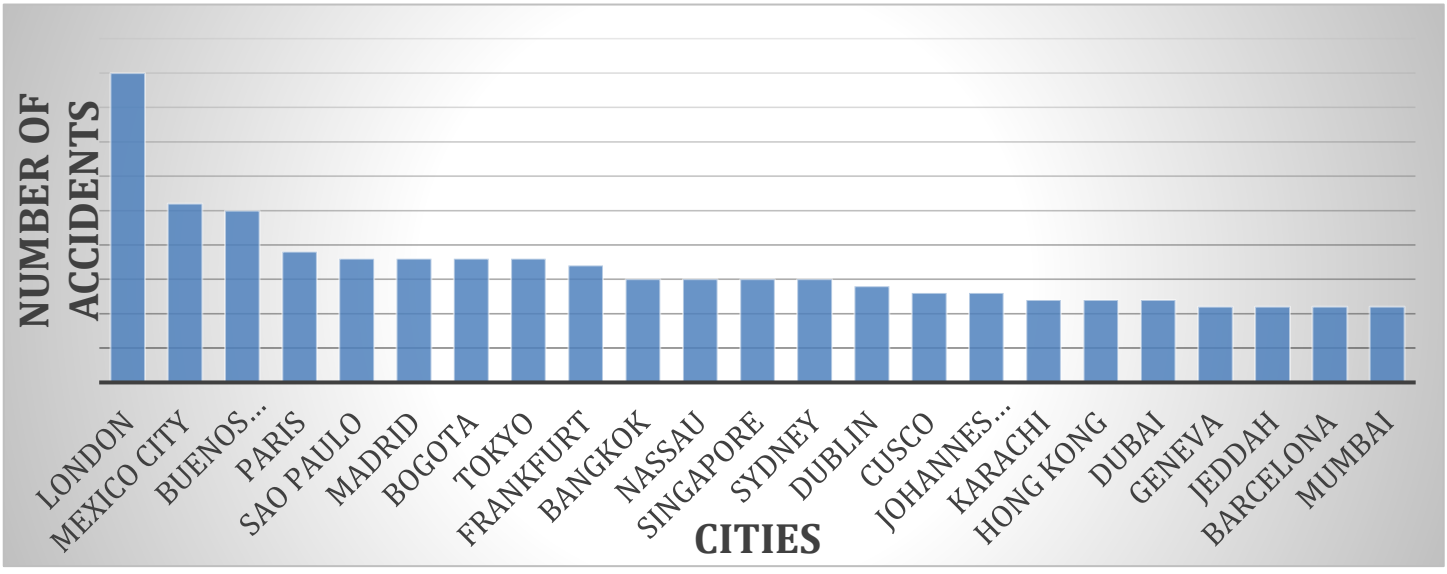
Q No 7: What cities (except USA) have observed the most accidents?

```
select count(*) as total_events, ev_city from events where ev_country <> 'USA'
group by ev_city order by total_events desc;
```

150 %

Results | Messages

	total_events	ev_city
1	45	London
2	26	Mexico City
3	25	Buenos Aires
4	19	Paris
5	18	Sao Paulo
6	18	Madrid
7	18	Bogota
8	18	Tokyo
9	17	Frankfurt
10	15	Bangkok
11	15	Nassau
12	15	Singapore
13	15	Sydney
14	14	Dublin
15	13	Cusco
16	13	Johannesburg
17	12	Karachi



London is the city with the highest air traffic accidents with the highest number of events reported outside the USA.

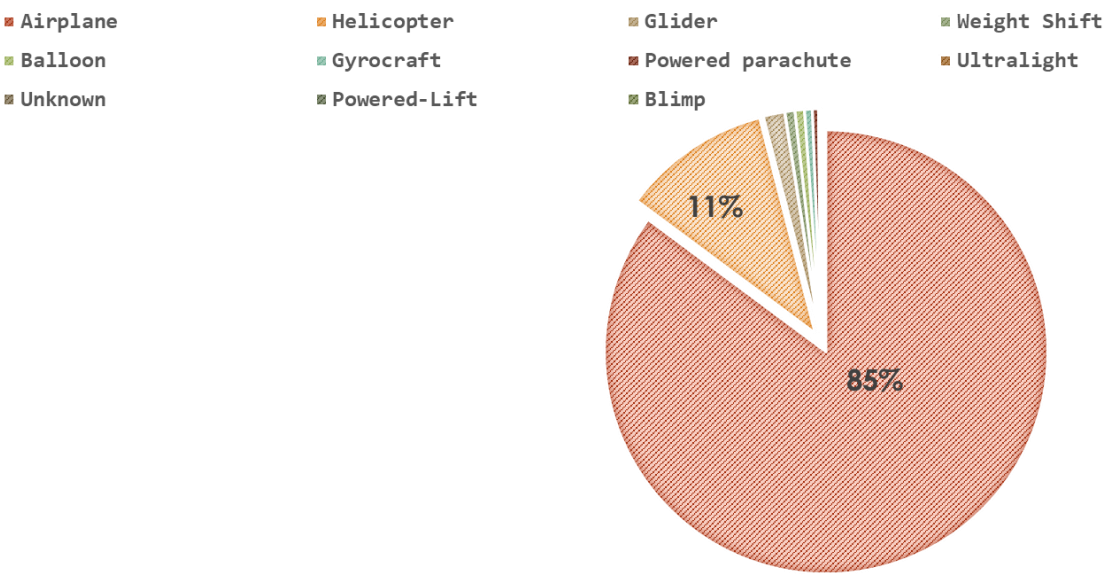
Q No 8: What is the ratio of aircraft category for total number of accidents?

```
select count(*) as aircraft_count, aircraft.acft_category
from events join aircraft on events.ev_id = aircraft.ev_id
where aircraft.acft_category <> ''
group by acft_category order by aircraft_count desc;
```

150 %

Results | Messages

	aircraft_count	acft_category
1	18332	AIR
2	2329	HELI
3	328	GLI
4	150	WSFT
5	142	BALL
6	118	GYRO
7	90	PPAR
8	25	ULTR
9	10	UNK
10	2	PLFT
11	1	BLIM
12	1	RCKT



85% of accidents were reported in airplanes with 11% in Helicopters and the rest 4% are others.

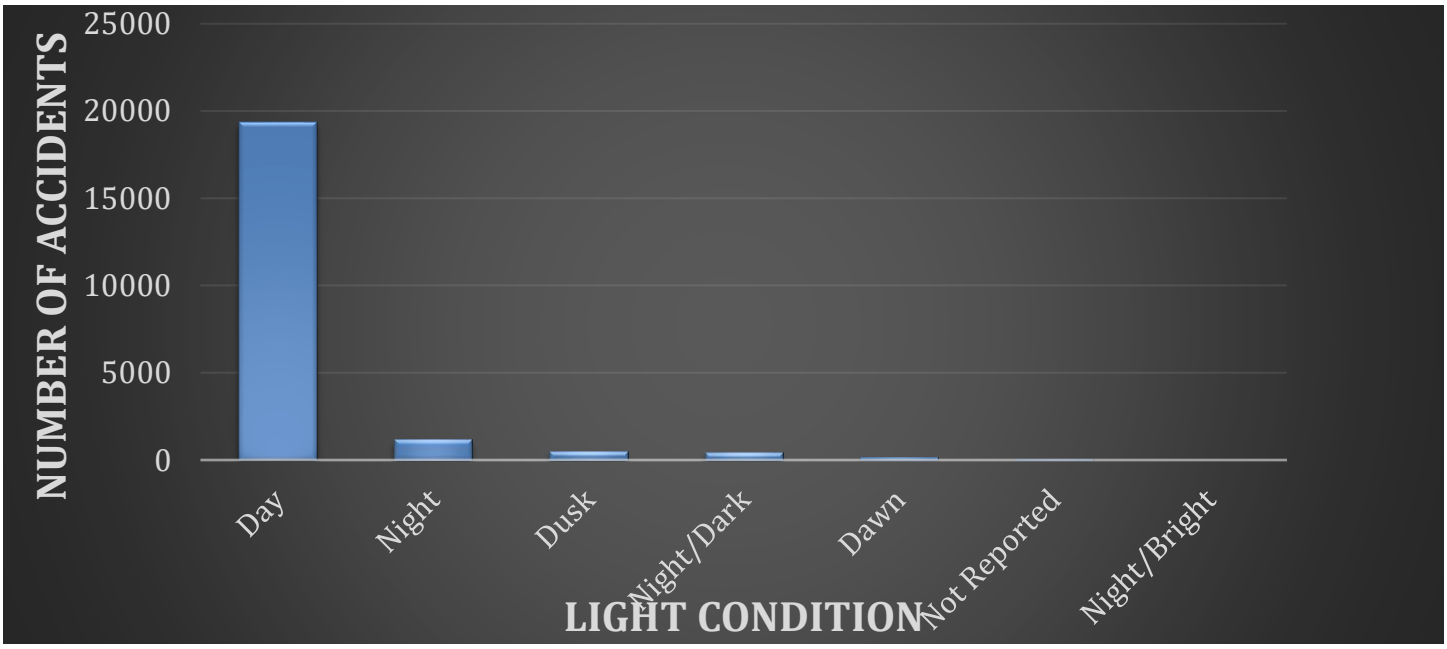
Q No 9: How are the accidents categorized in light conditions at the time of accidents?

```
select count(*) as total_events ,light_cond
from events where light_cond <> ''
GROUP BY light_cond ORDER BY total_events desc;
```

150 %

Results | Messages

	total_events	light_cond
1	19363	DAYL
2	1157	NITE
3	506	DUSK
4	431	NDRK
5	150	DAWN
6	85	NR
7	29	NBRT



Above **19000 accidents** occurred in the daylight

Q No 10: Which city has observed the most number of mid-air collisions of aircraft?

```
select count(distinct(events.ev_id)) as event_count, events.ev_city
from events
join aircraft on events.ev_id = aircraft.ev_id
join engines on engines.ev_id = events.ev_id
where acft_category = 'AIR' and
events.mid_air = 1
group by events.ev_city order by event_count desc;
```

150 %

Results | Messages

	event_count	ev_city
1	3	Reno
2	2	Wasilla
3	2	Fairbanks
4	2	Talkeetna
5	1	Tempsford, United Kingdom
6	1	Toms River
7	1	Tucson
8	1	Tuscaloosa
9	1	Fairmont

```
select  sum(subquery1.event_count) from
(select count(distinct(events.ev_id)) as event_count, events.ev_city
from events
join aircraft on events.ev_id = aircraft.ev_id
join engines on engines.ev_id = events.ev_id
where aircraft.acft_category = 'AIR' and
events.mid_air = 1
--and events.ev_country <> 'USA'
group by events.ev_city ) as subquery1
where subquery1.event_count = 1;
```

%

Results | Messages

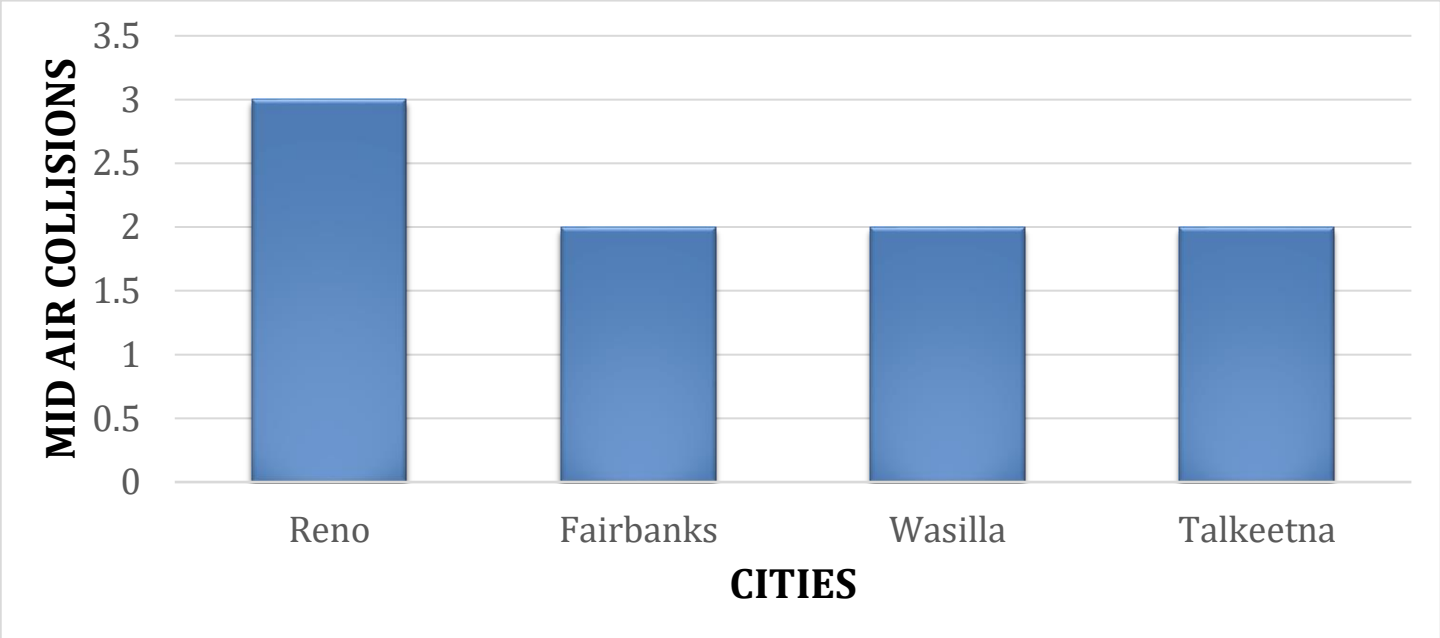
(No column name)
89

```sql
select  sum(subquery1.event_count) from
(select count(distinct(events.ev_id)) as event_count, events.ev_city
from events
join aircraft on events.ev_id = aircraft.ev_id
join engines on engines.ev_id = events.ev_id
where aircraft.acft_category = 'AIR' and
events.mid_air = 1
--and events.ev_country <> 'USA'
group by events.ev_city ) as subquery1
```

50 %

Results | Messages

	(No column name)
1	98



Out of **98 mid-air collisions**, **Reno** City observed 3 cases of mid-air collisions with **Fairbanks**, **Talkeetna**, and **Wasilla** with 2 cases each. And **89** other cities which once experienced mid-air collisions

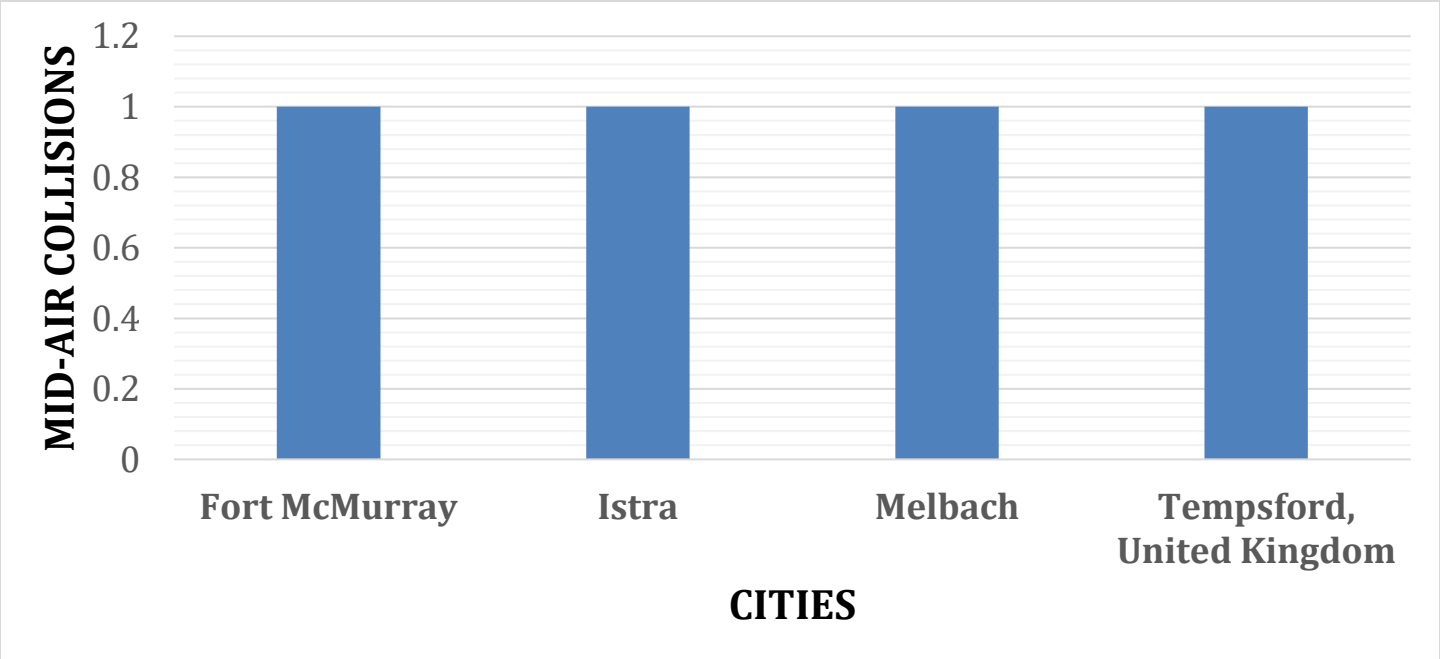
Q No 11: Which city (except USA) has observed most number of mid-air collisions of aircrafts?

```
select count(distinct(events.ev_id)) as event_count, events.ev_city
from events
join aircraft on events.ev_id = aircraft.ev_id
join engines on engines.ev_id = events.ev_id
where aircraft.acft_category = 'AIR' and
events.mid_air = 1
and events.ev_country <> 'USA'
group by events.ev_city order by event_count desc
```

150 %

Results | Messages

	event_count	ev_city
1	1	Fort McMurray
2	1	Istra
3	1	Melbach
4	1	Tempsford, United Kingdom



Each four cities observed a mid-air collision once since 2018.

Q No 12: For those aircraft with reciprocating (piston) engines, what was the fuel injection in which the explosion occurred inside the plane.

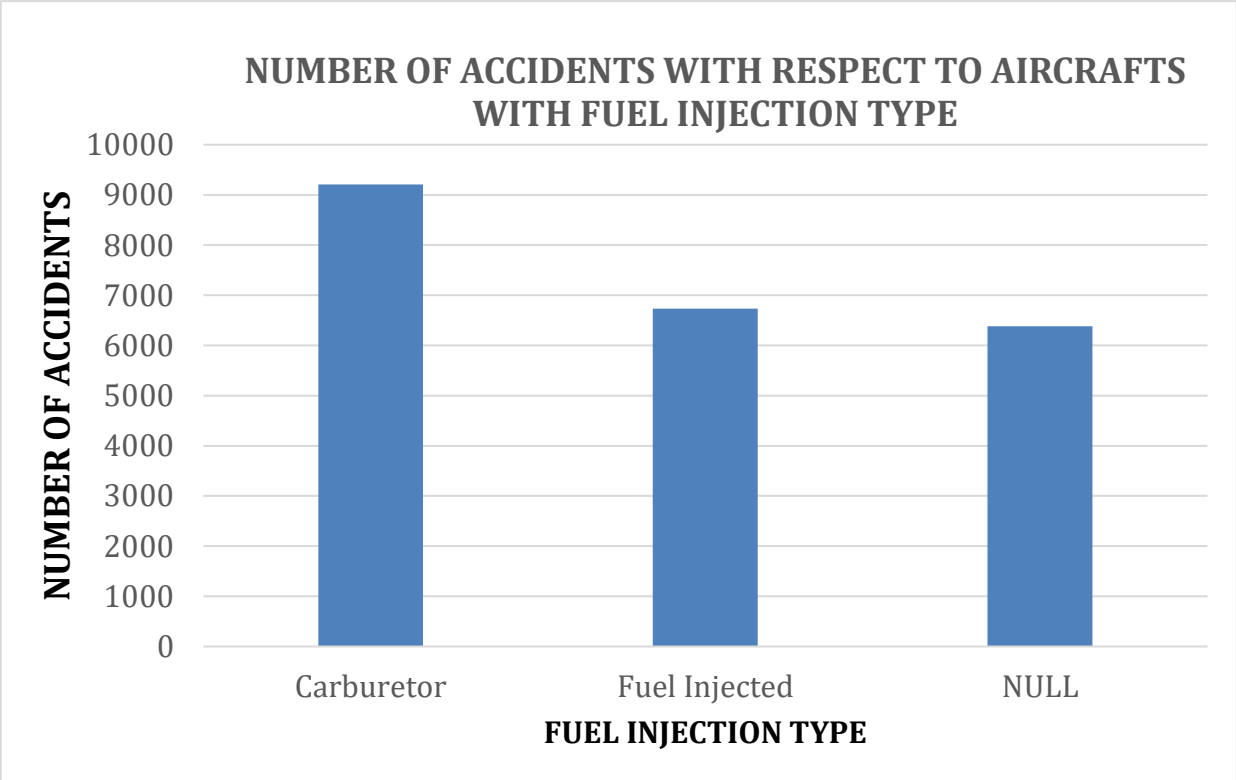
PART A:

```
select count(events.ev_id) as event_count, engines.carb_fuel_injection
from events
join aircraft on events.ev_id = aircraft.ev_id
join engines on events.ev_id = engines.ev_id
--where aircraft.acft_expl = 'IFLT'
group by engines.carb_fuel_injection order by event_count desc;
```

150 %

Results | Messages

	event_count	carb_fuel_injection
1	9210	CARB
2	6736	FINJ
3	6382	NULL



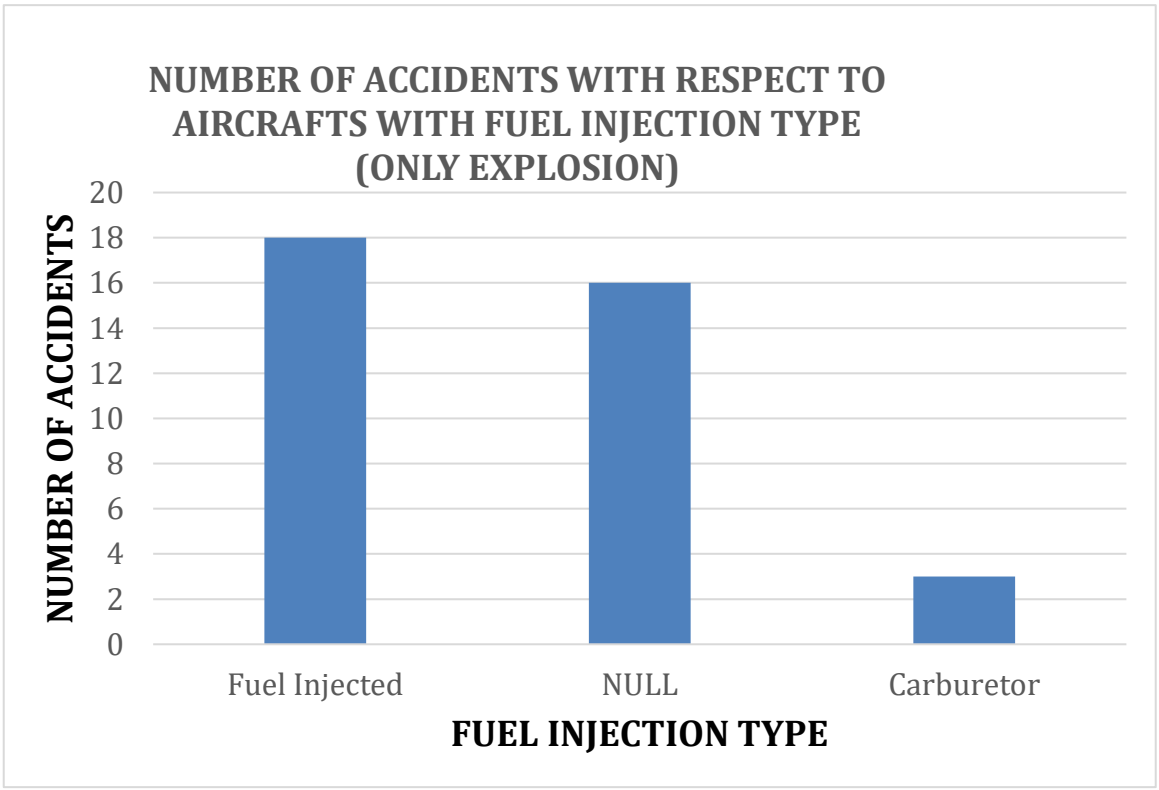
PART B:

```
select count(events.ev_id) as event_count, engines.carb_fuel_injection
from events
join aircraft on events.ev_id = aircraft.ev_id
join engines on events.ev_id = engines.ev_id
where aircraft.acft_expl in ('IFLT','BOTH')
group by engines.carb_fuel_injection order by event_count desc;
```

150 %

Results | Messages

	event_count	carb_fuel_injection
1	18	FINJ
2	16	NULL
3	3	CARB



For all the events the Carburetor engine is showing the bigger graph. But with the accidents

where the explosion took place the engines with fuel-injection system is way above the other.

Recommendations:

- The air spaces of cities with the most accidents reported, especially with the mid-air collisions should be monitored more strictly.
- More experienced pilots should be given priority in sensitive areas in the peak months of the year.
- The daytime routine should be monitored with more focus in the future.
- The planes with engines having fuel injected type must be given proper maintenance after each time interval to avoid engine explosion.

Group Members:

- Arsalan Ahmad
- Fazal Hannan
- Arshian Bashir