AIRLINES DATA ANLYTICS WITH QLIK CLOUD

SUBMMITED BY

PADALA RAHUL

1.INTRODUCTION

1.1 OVERVIEW:

In order to unleash valuable insights from airline data, we explored the potential of Qlik Cloud which is a cloud-based data visualization platform. Our aim was to provide airlines with interactive dashboards that would enable their stakeholders make informed decisions by converting raw information into attractive visuals.

1.2 PURPOSE:

I had three main goals:

1. Definition of Key Indicators: I have found the most important measures that show the state and efficiency of airline work. Such KPIs could include the percentage of punctual flights, client characteristics or network connections.

2. Creation of effective data visualizations: My task was to turn recognized KPIs into simple visual charts. For this purpose, we used graphs, maps and tables which could represent difficult information in an easily understandable way.

3. Development of interactive panels: My objective lay in establishment of interactive Dashboards serving as a basic point for exploration of air carrier statistics. These Dashboards enable one to filter, drill down into and analyze facts from different angles.

1.3 TECHNICAL ARCHITECHTURE:

The foundation of this project relied on the following technologies:

Data Source: Airline database (The data source is provided by Smart bridge platform)

Data Visualization Tool: Qlik Cloud

Development Tools: Qlik Cloud, Insight Advisor

2.DEFINE PROBLEM / PROBLEM UNDERSTANING

2.1 SPECIFY THE BUSINESS PROBLEM

I aim to look at the main features of airline operations.determining coverage by geography as well as quantifying delays of flights alongside their performance on time and cancellations. It is important also to establish how many people delays and cancellations affect so that we can know their importance. The knowledge acquired from this analysis will be used for our airlines service improvement in terms of operational efficiency, reliability ultimately customer satisfaction which will inform strategic decision making towards this end.

2.2 BUSINESS REQUIREMENTS

In order to meet the objectives of Airlines Database Analytics project we shall put in place mechanisms for collection and analysis of customer information, map out geographical coverage as well as watch flight performance among other things delays, on-time percentages and cancellations. We will also look at how these disturbances affect travelers so as to know their gravity. This is to be achieved through setting up standards by which performance can be measured and coming up with different views for both immediate and past records with an intention of enhancing efficiency in operations besides satisfying clients' needs at all times. Moreover we are going to collect opinions from our customers showing us where we need to improve on while at the same time making sure that their data is secured and private. Strategic decisions will then be made based upon insights obtained which may include but not limited to service improvement among others.

2.3 LITERATURE SURVEY

Research in airline operations emphasizes the importance of comprehensive data collection and analysis to understand travel patterns and optimize services. Mapping geographical coverage is crucial for identifying new market opportunities and improving route networks. Monitoring flight performance, particularly delays and cancellations, enhances operational efficiency and customer satisfaction. Assessing the impact of these disruptions on passengers is vital for developing effective mitigation strategies. Key performance indicators (KPIs) like delay times and cancellation rates provide insights into operational

efficiency. Collecting customer feedback is essential for identifying service improvement areas. Predictive analytics support strategic decision-making by forecasting trends and challenges. Robust data security measures are crucial to protect passenger information and maintain trust. Together, these elements contribute to a more efficient, reliable, and customer-centric airline operation.

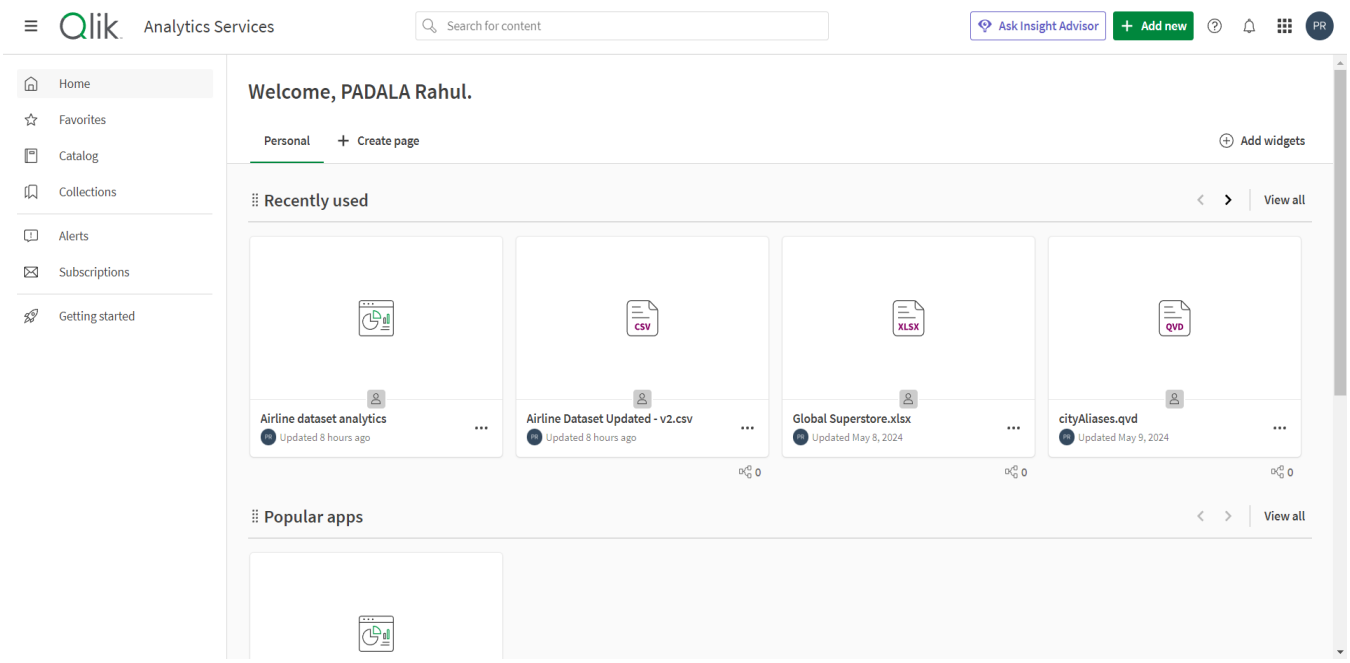
3.DATA COLLECTION

3.1 COLLECT THE DATASET

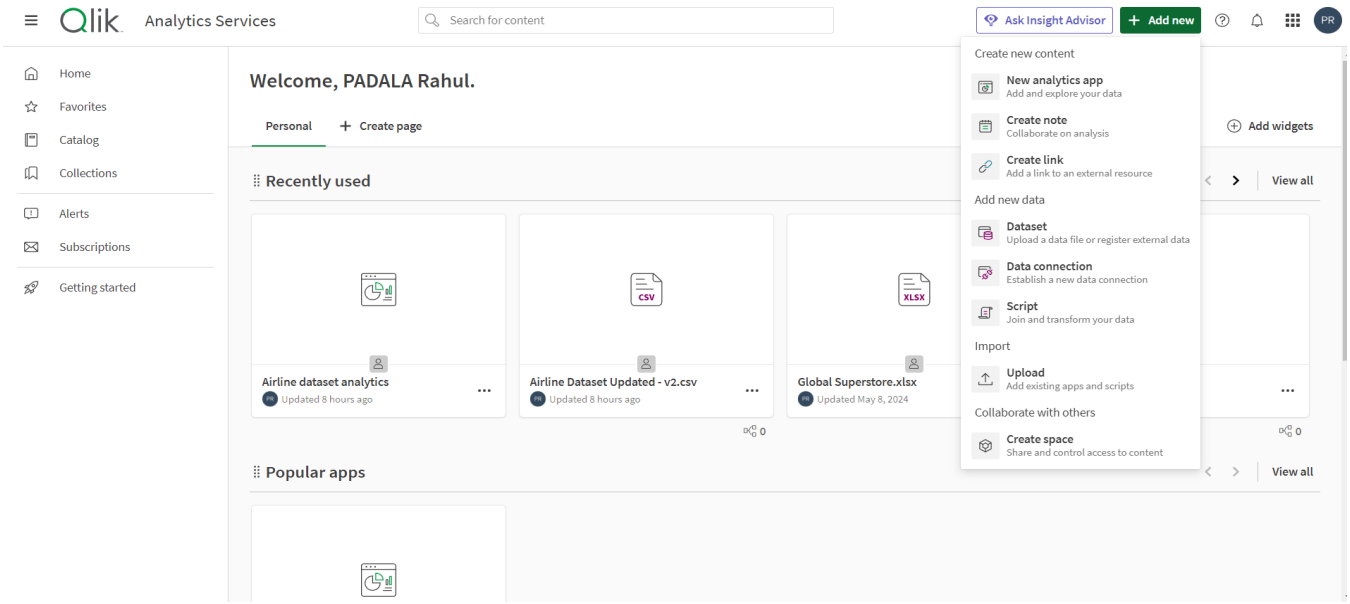
The dataset is collected from the smartbridge platform.

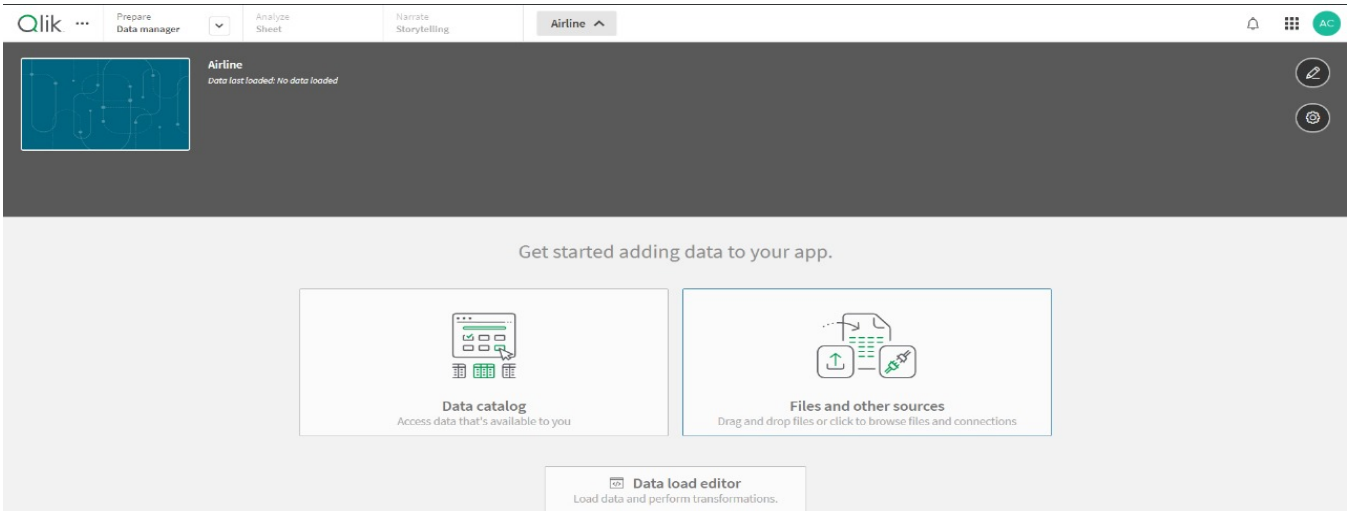
3.2 CONNECT DATA WITH QLIK SENSE

1.Once we logined into qlik cloud we will be getting interface shown below



2.On the interface we will be find a button called Add new click on that we will get a drop down in that drop down we will find a button called new analytics app then click on that after that it will be showing upload dataset option then select file from our device..





after the above process we data connected will with qlik sense.

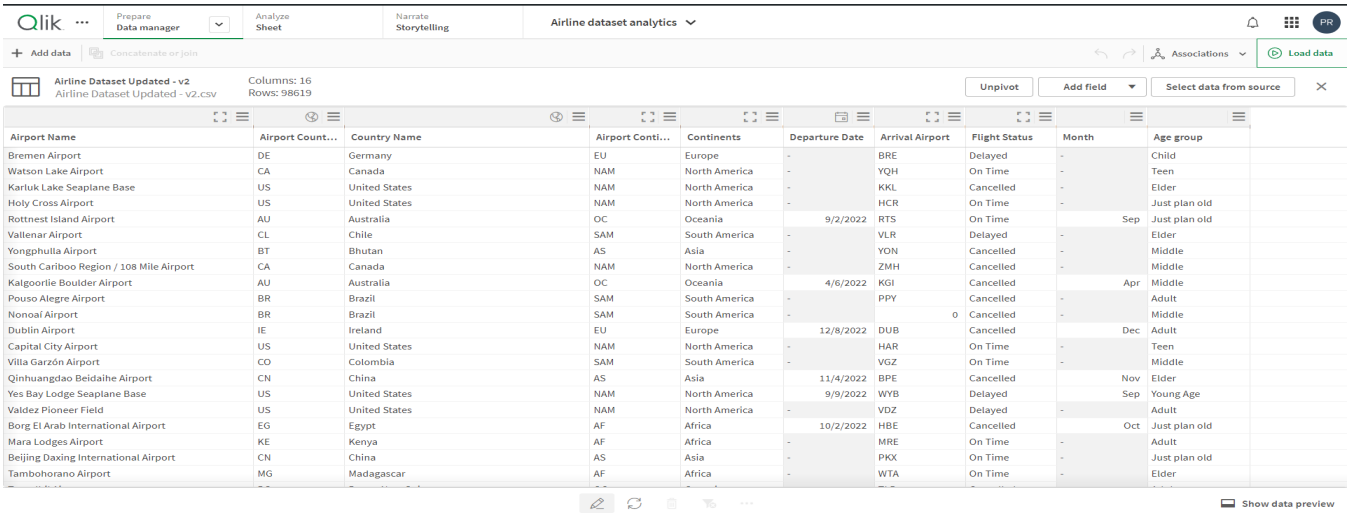
4. DATA PREPARATION

4.1 PREPARE THE DATA FOR VISUALIZATION

After connecting dataset with qlik we need to clean and preprocess the data based on our reqirement.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Passenger	First Name	Last Name	Gender	Age	Nationality	Airport Na	Airport Co	Country Na	Airport Co	Continents	Departure	Arrival Airp	Pilot Name	Flight Stat
2	ABVWIg	Edithe	Leggis	Female	62	Japan	Coldfoot A	US	United Sta	NAM	North Ame	6/28/2022	CXF	Fransisco I	On Time
3	jkXXAX	Elwood	Catt	Male	62	Nicaragua	Kugluktuk ,	CA	Canada	NAM	North Ame	12/26/202	YCO	Marla Pars	On Time
4	CdUz2g	Darby	Felgate	Male	67	Russia	Grenoble-I	FR	France	EU	Europe	1/18/2022	GNB	Rhonda Ar	On Time
5	BRS38V	Dominica	Pyle	Female	71	China	Ottawa / C	CA	Canada	NAM	North Ame	9/16/2022	YND	Kacie Com	Delayed
6	9kvTLo	Bay	Pencost	Male	21	China	Gillespie Fi	US	United Sta	NAM	North Ame	2/25/2022	SEE	Ebonee Tr	On Time
7	nMJKVh	Lora	Durbann	Female	55	Brazil	Coronel Hı	BR	Brazil	SAM	South Ame	########	LEC	Inglis Dolle	On Time
8	8IPFPE	Rand	Bram	Male	73	Ivory Coas	Duxford Ai	GB	United Kin	EU	Europe	10/30/202	QFO	Stanislas T	Cancelled
9	pqixbY	Perceval	Dallosso	Male	36	Vietnam	Maestro W	BR	Brazil	SAM	South Ame	########	STM	Sharyl East	Cancelled
10	QNAs2R	Aleda	Pigram	Female	35	Palestiniar	Venice Ma	IT	Italy	EU	Europe	8/20/2022	VCE	Daryn Barı	On Time
11	3jmudz	Burlie	Schustl	Male	13	Thailand	Vermilion ,	CA	Canada	NAM	North Ame	########	YVG	Alameda C	On Time
12	2P41gZ	Porty	Jori	Male	39	Tunisia	Nuevo Cas	MX	Mexico	NAM	North Ame	5/27/2022	NCG	Rasia Fidel	Cancelled
13	sBf524	Briant	De La Hay	Male	71	Russia	Ruben Can	PA	Panama	NAM	North Ame	########	SYP	Alina Flool	Delayed
14	PlwJZT	Kalie	Scoble	Female	47	Sweden	Loralai Air	PK	Pakistan	AS	Asia	3/19/2022	LRG	Madelena	Delayed
15	iU75x3	Catriona	Beaument	Female	77	Russia	Cudal Airp	AU	Australia	OC	Oceania	3/24/2022	CUG	Margie Be	Delayed
16	GUta6R	Amberly	Handling	Female	32	China	Farmingto	US	United Sta	NAM	North Ame	########	FAM	Lothaire E	Delayed
17	8qA80a	Dyna	De'Vere - I	Female	22	China	Oudtshoor	ZA	South Afric	AF	Africa	7/18/2022	OUH	Neila Gier	Cancelled
18	2haCDu	Janella	Hardaker	Female	28	Colombia	Zaraza Air	VE	Venezuela	SAM	South Ame	9/23/2022	ZRZ	Shaylynn A	On Time
19	WiVl8o	Alvin	Wenzel	Male	12	Greece	Enshi Airpo	CN	China	AS	Asia	3/29/2022	ENH	Alfie Mac	Delayed
20	e0H5LI	Jerrine	Peeters	Female	87	Philippines	Thompson	CA	Canada	NAM	North Ame	########	YTH	Chandra D	Cancelled
21	nL8kyD	Warner	Driutti	Male	62	China	Guilin Lian	CN	China	AS	Asia	########	KWL	Marita Ho	Cancelled
22	9iT79e	Paige	Hayhow	Male	24	Sweden	Crested Bu	US	United Sta	NAM	North Ame	2/19/2022	CSE	Clyde Win	Delayed
23	kEARqP	Dorisa	Skill	Female	19	Ukraine	St Augustir	CA	Canada	NAM	North Ame	########	YIF	Ilyse Bartl	Cancelled
24	dx3NWh	Bobbye	Patmore	Female	45	China	Port BergÃ	MG	Madagasc	AF	Africa	8/24/2022	WPB	Stella Pitth	On Time
25	Dpafly	Jayme	Dairton	Female	29	Indonesia	Hato Coro	CO	Colombia	SAM	South Ame	########	HTZ	Kaye Clew	Delayed
26	WV3aXu	Sayre	Stroyan	Male	87	Indonesia	Elkhart Mu	US	United Sta	NAM	North Ame	3/20/2022	EKI	Austine Cr	Cancelled
27	f27tyG	Joellyn	Stutter	Female	8	Croatia	Timbedra ,	MR	Mauritania	AF	Africa	2/19/2022	TMD	Janaya Ret	On Time

At the stage of cleaning and preprocessing I removed field pilot name and addded fields like month and age group which are use for our visualization. you can see in the below figure.

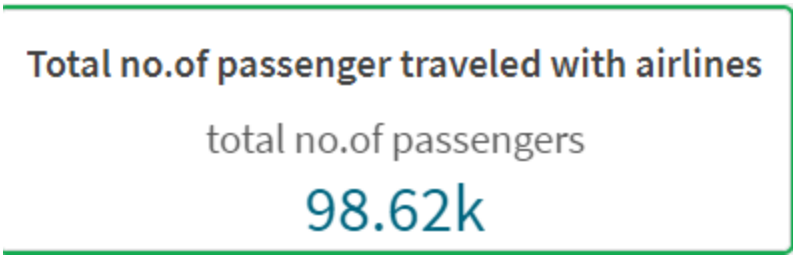


Airport Name	Airport Count...	Country Name	Airport Conti...	Continents	Departure Date	Arrival Airport	Flight Status	Month	Age group
Bremen Airport	DE	Germany	EU	Europe	-	BRE	Delayed	-	Child
Watson Lake Airport	CA	Canada	NAM	North America	-	YQH	On Time	-	Teen
Karluk Lake Seaplane Base	US	United States	NAM	North America	-	KKL	Cancelled	-	Elder
Holy Cross Airport	US	United States	NAM	North America	-	HCR	On Time	-	Just plan old
Rottnest Island Airport	AU	Australia	OC	Oceania	9/2/2022	RTS	On Time	Sep	Just plan old
Vallenar Airport	CL	Chile	SAM	South America	-	VLR	Delayed	-	Elder
Yongphulla Airport	BT	Bhutan	AS	Asia	-	YON	Cancelled	-	Middle
South Cariboo Region / 108 Mile Airport	CA	Canada	NAM	North America	-	ZMH	Cancelled	-	Middle
Kalgoorlie Boulder Airport	AU	Australia	OC	Oceania	4/6/2022	KGI	Cancelled	Apr	Middle
Pouso Alegre Airport	BR	Brazil	SAM	South America	-	PPY	Cancelled	-	Adult
Nonoaí Airport	BR	Brazil	SAM	South America	-	0	Cancelled	-	Middle
Dublin Airport	IE	Ireland	EU	Europe	12/8/2022	DUB	Cancelled	Dec	Adult
Capital City Airport	US	United States	NAM	North America	-	HAR	On Time	-	Teen
Villa Garzón Airport	CO	Colombia	SAM	South America	-	VGZ	On Time	-	Middle
Qinhuangdao Beidaihe Airport	CN	China	AS	Asia	11/4/2022	BPE	Cancelled	Nov	Elder
Yes Bay Lodge Seaplane Base	US	United States	NAM	North America	9/9/2022	WYB	Delayed	Sep	Young Age
Valdez Pioneer Field	US	United States	NAM	North America	-	VDZ	Delayed	-	Adult
Borg El Arab International Airport	EG	Egypt	AF	Africa	10/2/2022	HBE	Cancelled	Oct	Just plan old
Mara Lodges Airport	KE	Kenya	AF	Africa	-	MRE	On Time	-	Adult
Beijing Daxing International Airport	CN	China	AS	Asia	-	PKX	On Time	-	Just plan old
Tambohorano Airport	MG	Madagascar	AF	Africa	-	WTA	On Time	-	Elder

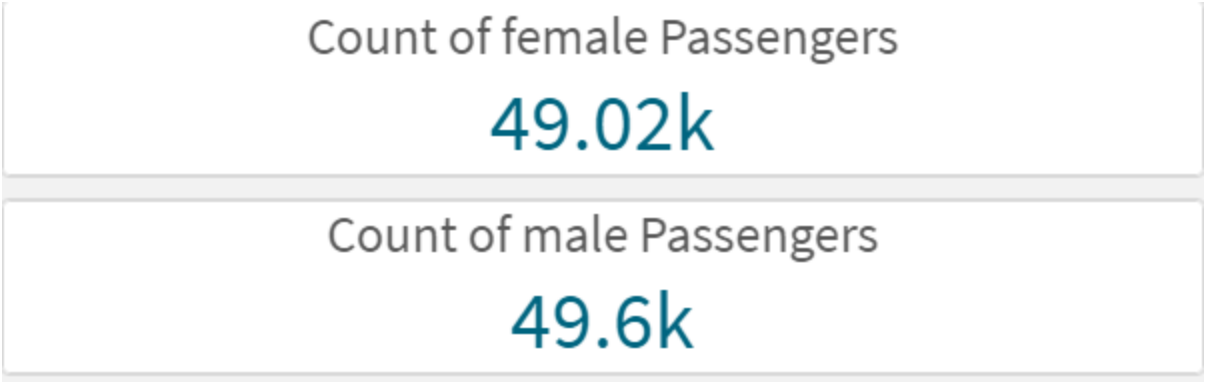
5. DATA VISUALIZATION

5.1 VISUALIZATION

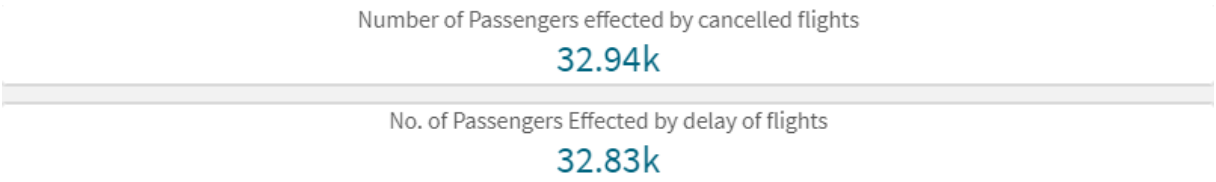
5.1.1



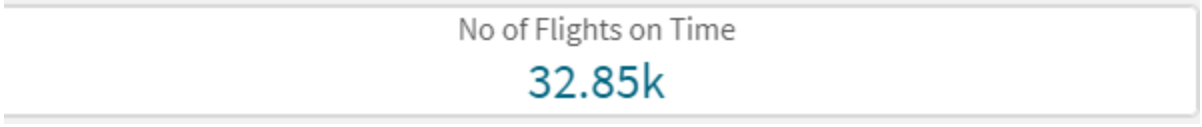
5.1.2



5.1.3



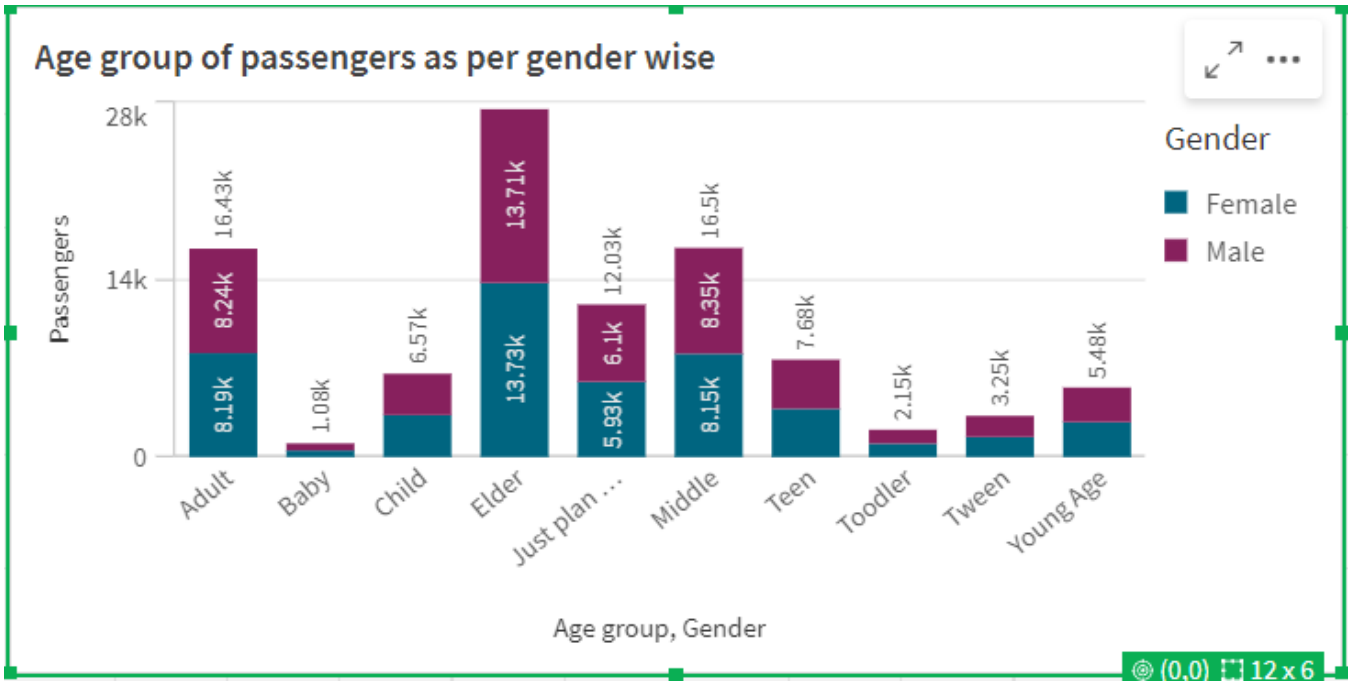
5.1.4



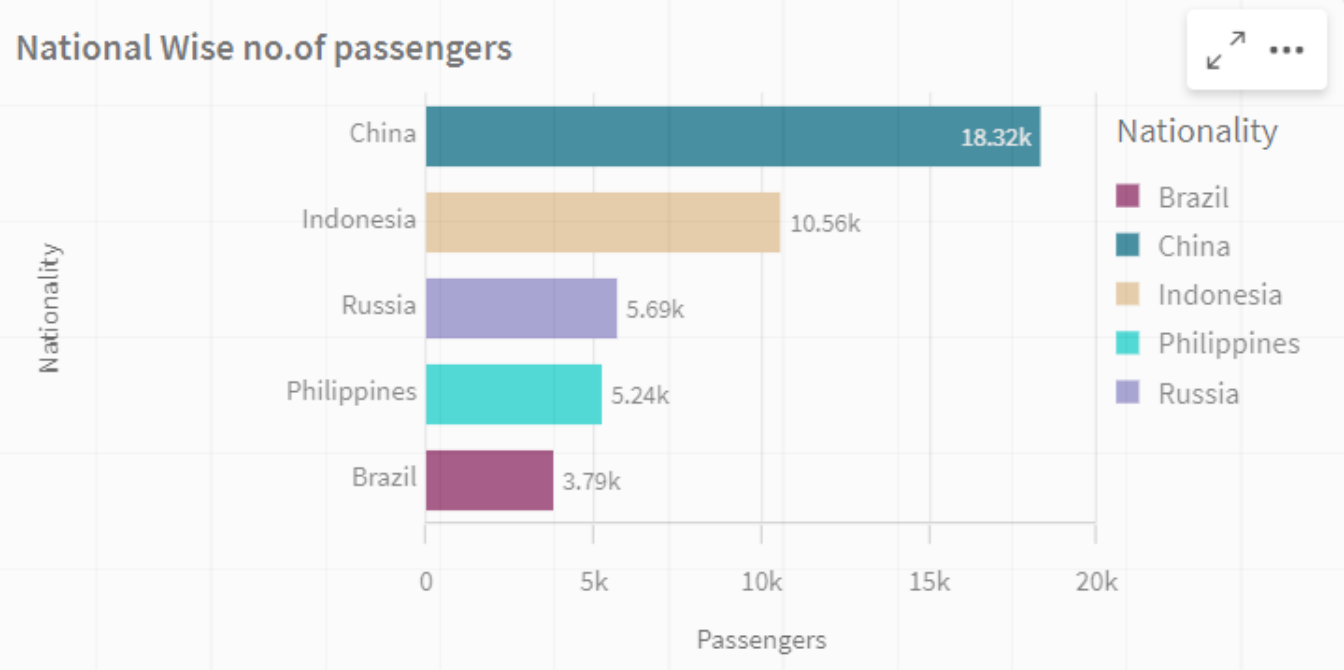
5.1.5


Top 5-No of Passengers travelled- Month Wise
4k
3k
2k
1k
0
Passengers
3.31k
3.31k
3.28k
3.26k
3.26k
May
Aug
Jul
Jun
Feb
Month
Month
Feb
May
Jun
Jul
Aug

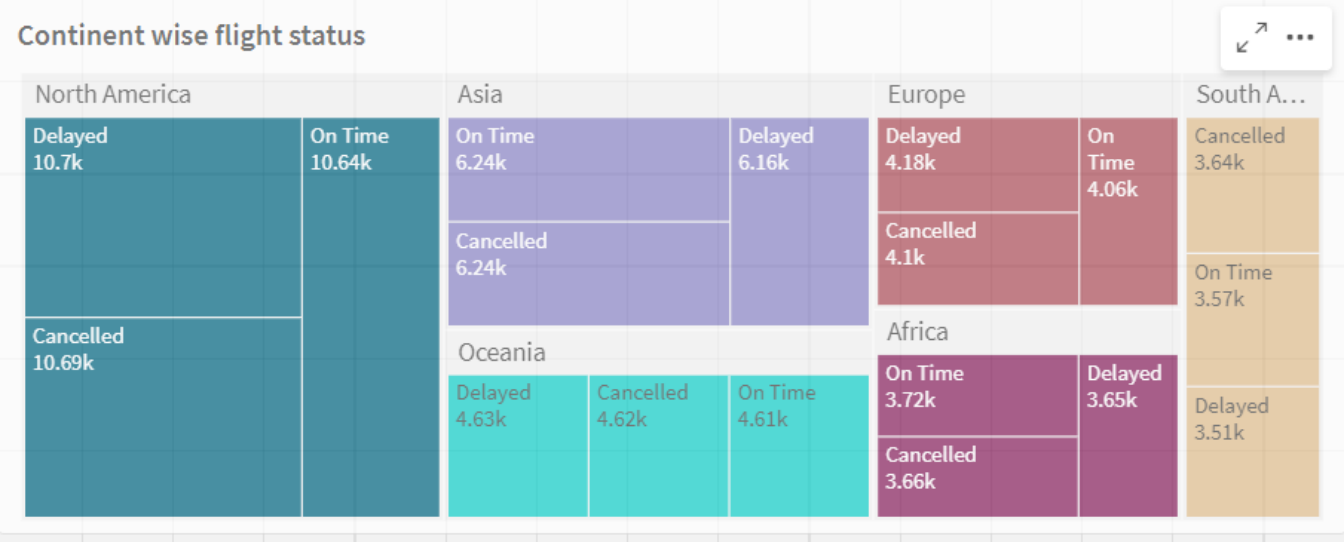
5.1.6


Age group of passengers as per gender wise
28k
14k
0
Passengers
16.43k
8.24k
8.19k
1.08k
6.57k
13.71k
13.73k
12.03k
6.1k
5.93k
16.5k
8.35k
8.15k
7.68k
2.15k
3.25k
5.48k
Adult
Baby
Child
Elder
Just plan ...
Middle
Teen
Toodler
Tween
Young Age
Age group, Gender
Gender
Female
Male
(0,0) 12 x 6

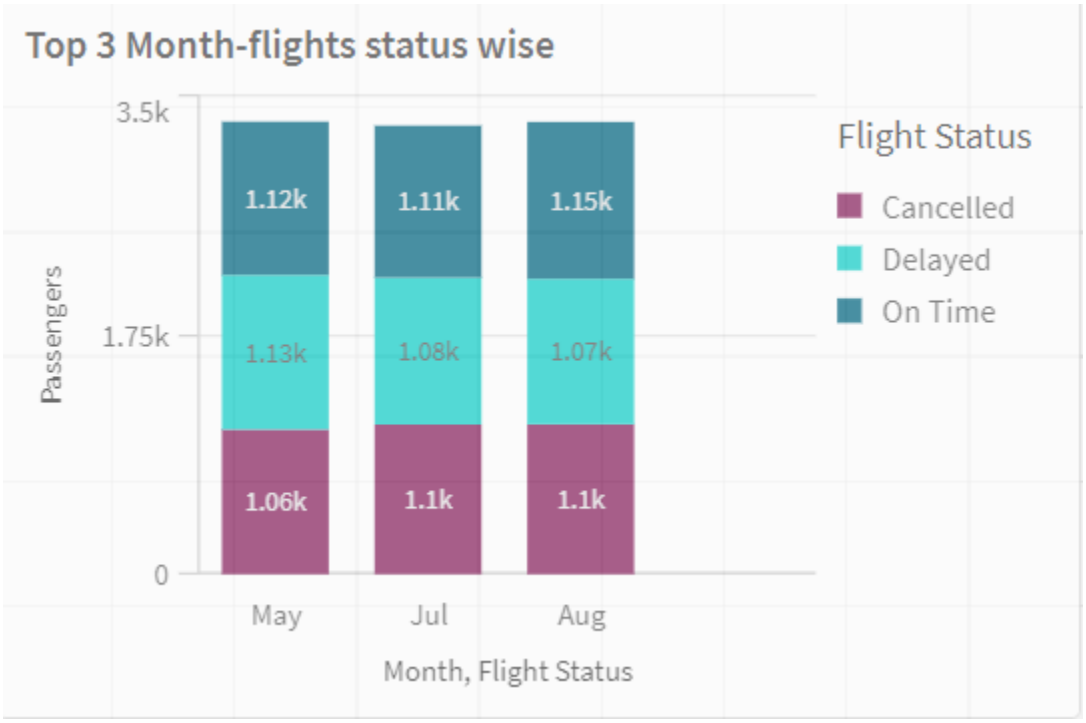
5.1.7


National Wise no.of passengers
China
Indonesia
Russia
Philippines
Brazil
18.32k
10.56k
5.69k
5.24k
3.79k
Nationality
0
5k
10k
15k
20k
Passengers
Nationality
Brazil
China
Indonesia
Philippines
Russia

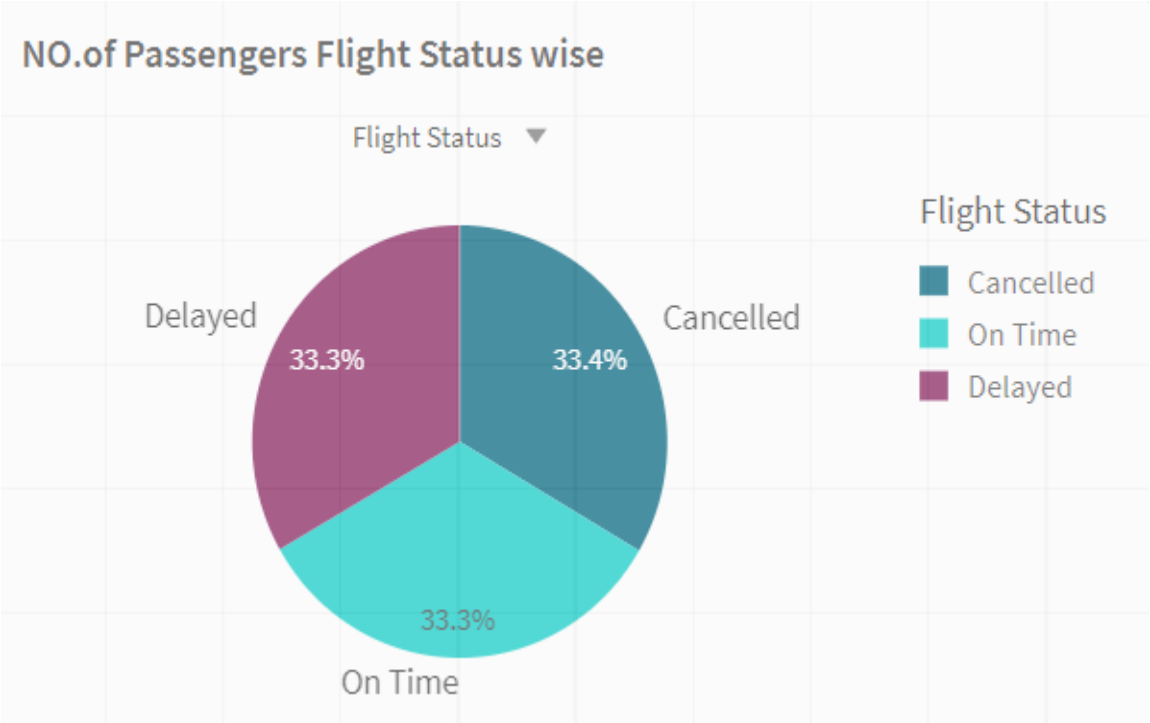
5.1.8


Continent wise flight status
North America
Delayed 10.7k
On Time 10.64k
Cancelled 10.69k
Asia
On Time 6.24k
Delayed 6.16k
Cancelled 6.24k
Oceania
Delayed 4.63k
Cancelled 4.62k
On Time 4.61k
Europe
Delayed 4.18k
On Time 4.06k
Cancelled 4.1k
Africa
On Time 3.72k
Delayed 3.65k
Cancelled 3.66k
South A...
Cancelled 3.64k
On Time 3.57k
Delayed 3.51k

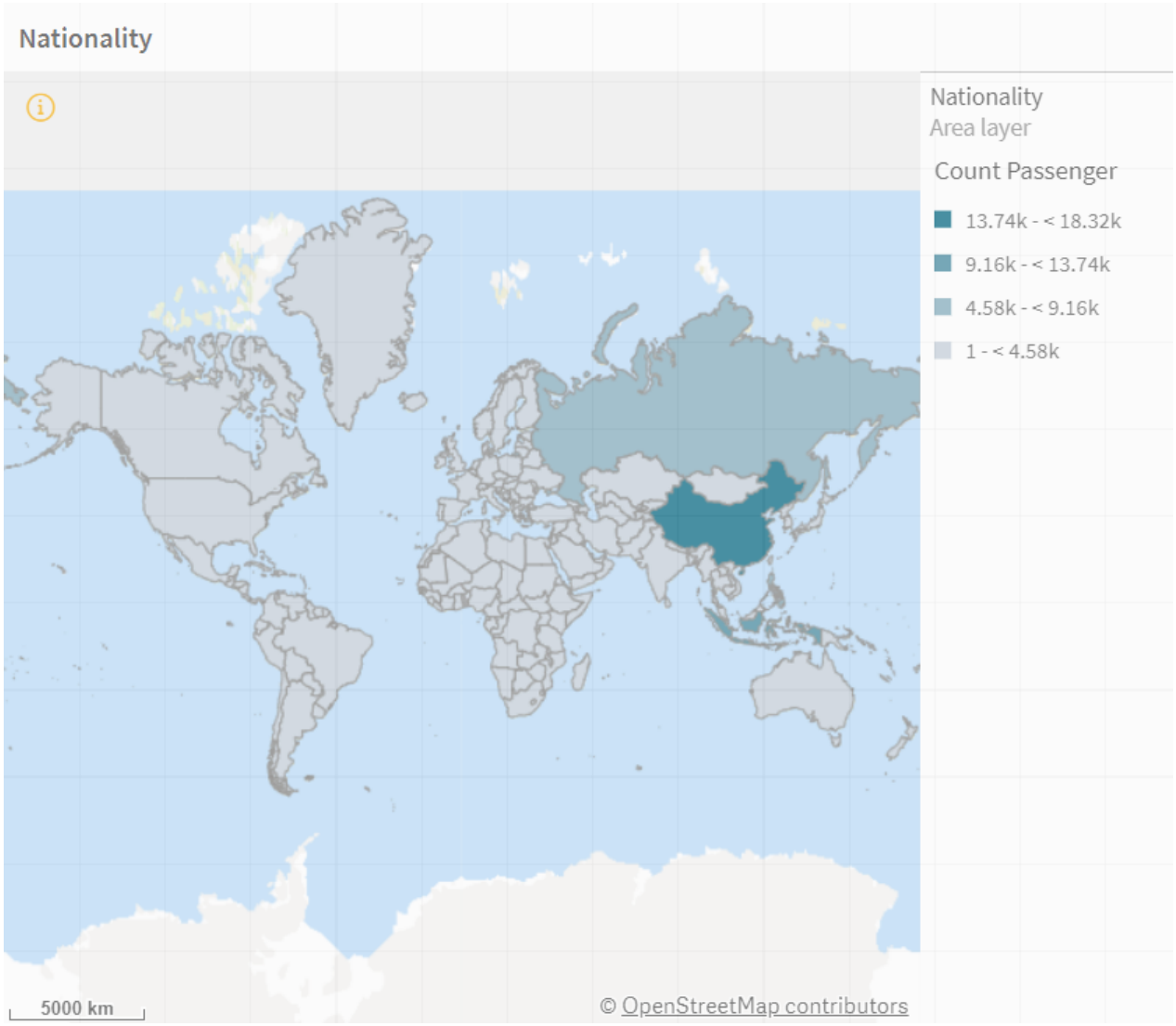
5.1.9


Top 3 Month-flights status wise
3.5k
1.75k
0
Passengers
1.12k
1.11k
1.15k
1.13k
1.08k
1.07k
1.06k
1.1k
1.1k
May
Jul
Aug
Month, Flight Status
Flight Status
Cancelled
Delayed
On Time

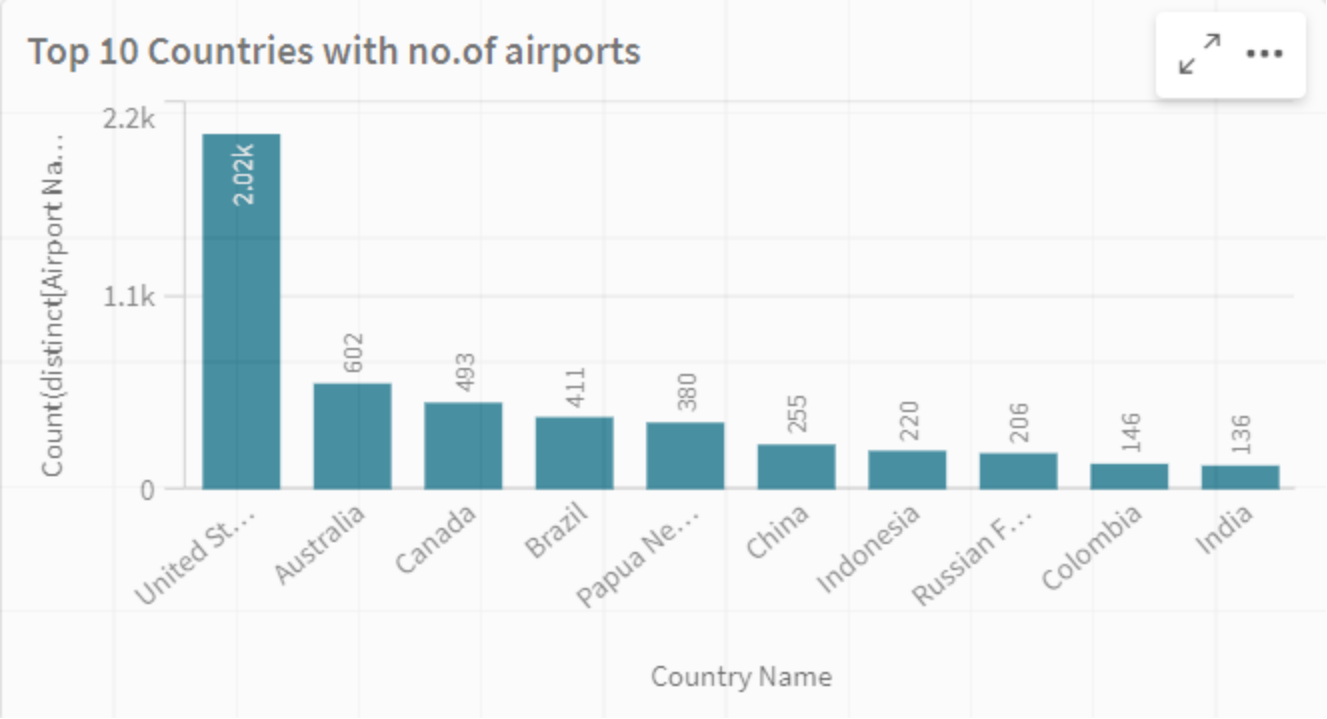
5.1.10


NO.of Passengers Flight Status wise
Flight Status
Delayed
Cancelled
33.3%
33.4%
33.3%
On Time
Flight Status
Cancelled
On Time
Delayed

5.1.11


Nationality
Nationality
Area layer
Count Passenger
13.74k - < 18.32k
9.16k - < 13.74k
4.58k - < 9.16k
1 - < 4.58k
5000 km
© OpenStreetMap contributors

5.1.12

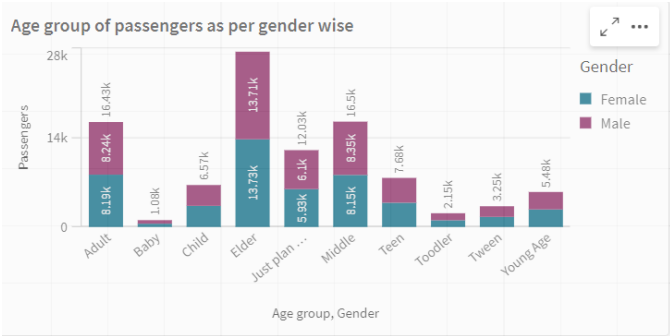

Top 10 Countries with no.of airports
Count(distinct[Airport Na...
2.2k
1.1k
0
2.02k
602
493
411
380
255
220
206
146
136
United St...
Australia
Canada
Brazil
Papua Ne...
China
Indonesia
Russian F...
Colombia
India
Country Name

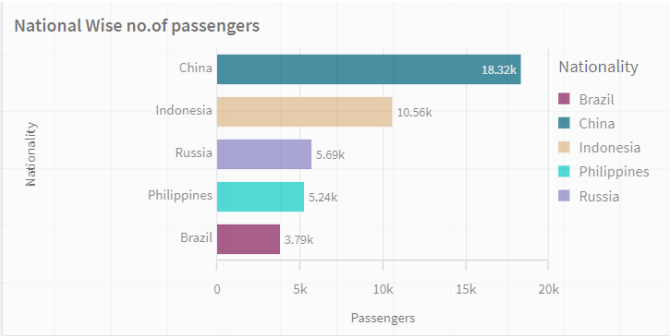
5.1.13

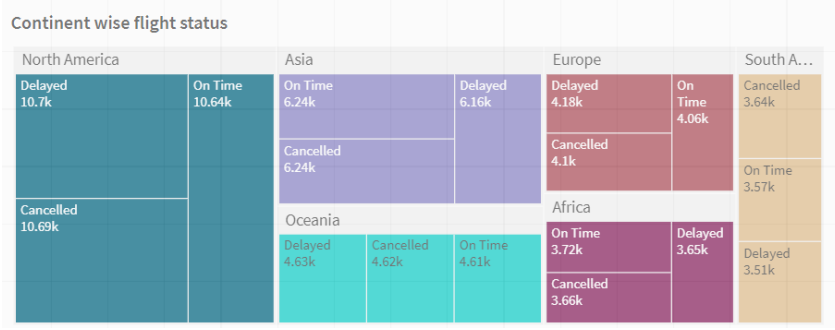

Fight Details
Count([Flight Status])
60
40
20
0
36
13
10
13
31
11
9
11
43
17
15
11
35
15
14
38
9
17
12
Böblinge...
San Fern...
San Pedr...
Santa An...
Santa Mar...
Airport Name, Flight Status
Flight Status
Cancelled
Delayed
On Time

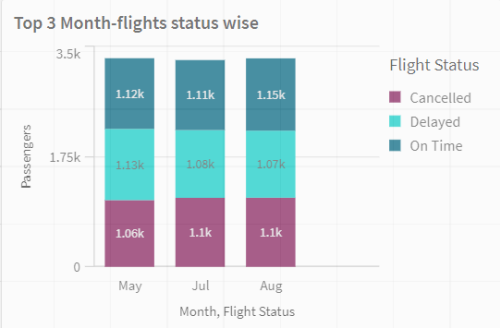
6. DASHBOARD

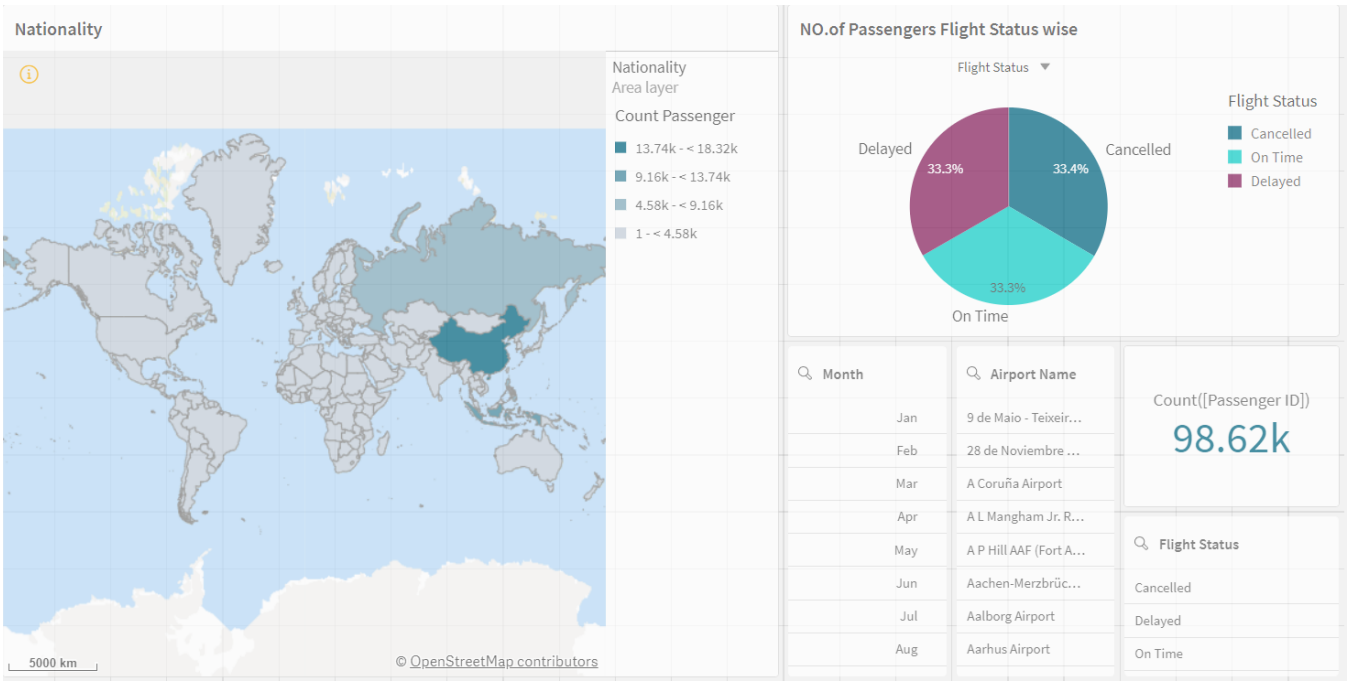
6.1 RESPONSIVE AND DESIGN OF DASHBOARD



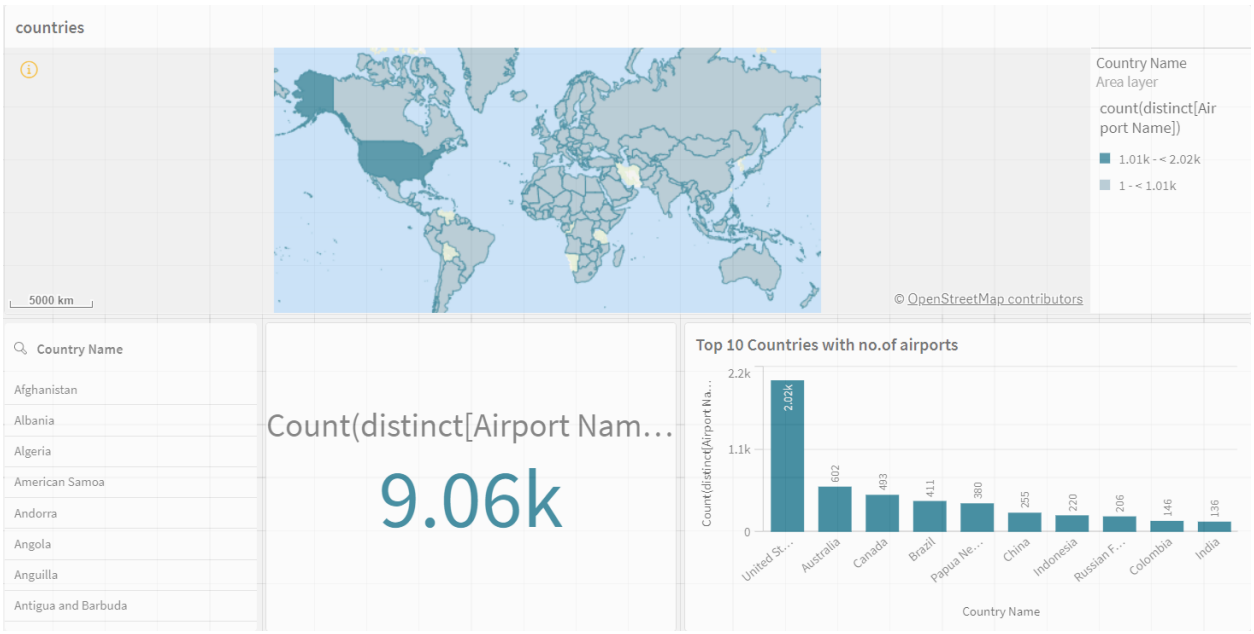




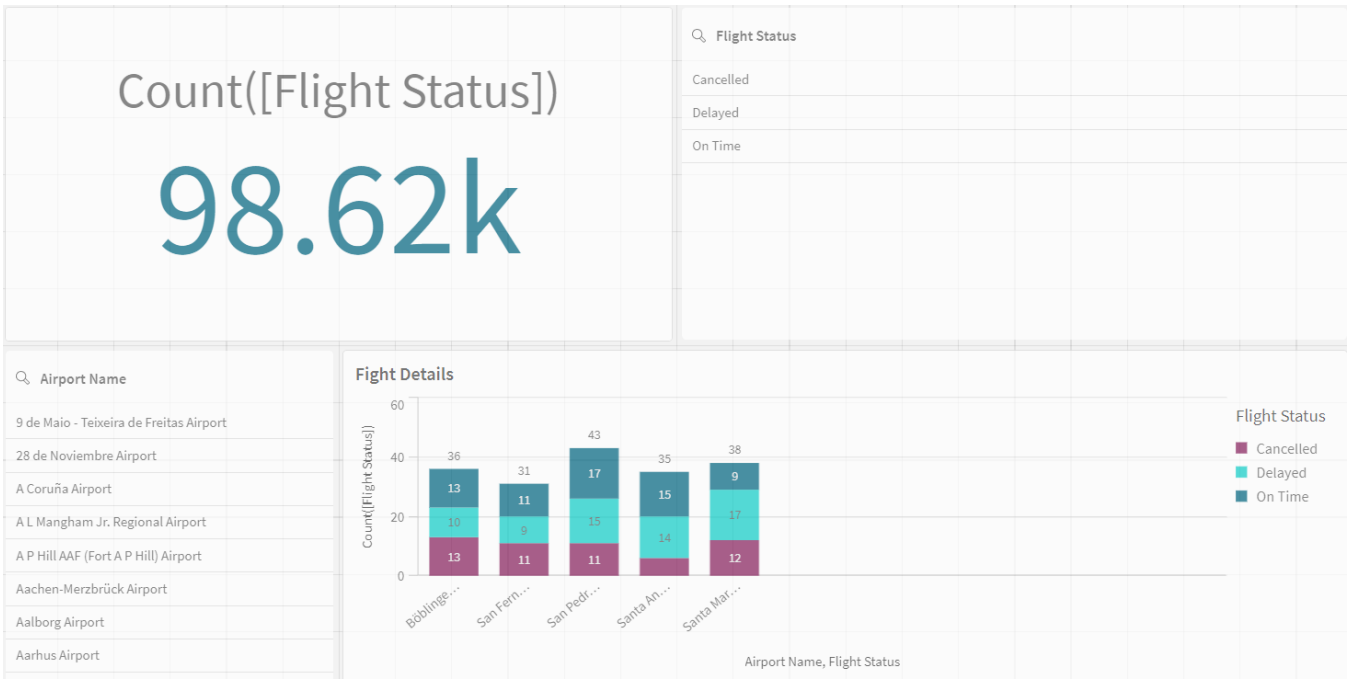




The above figure or Dashboard will changes Based on the applied filters the map and pie chart will be changing. in nationality we getting count of passengers based on groping nationality and pie chart is showing the no.of passengers flight status wise.



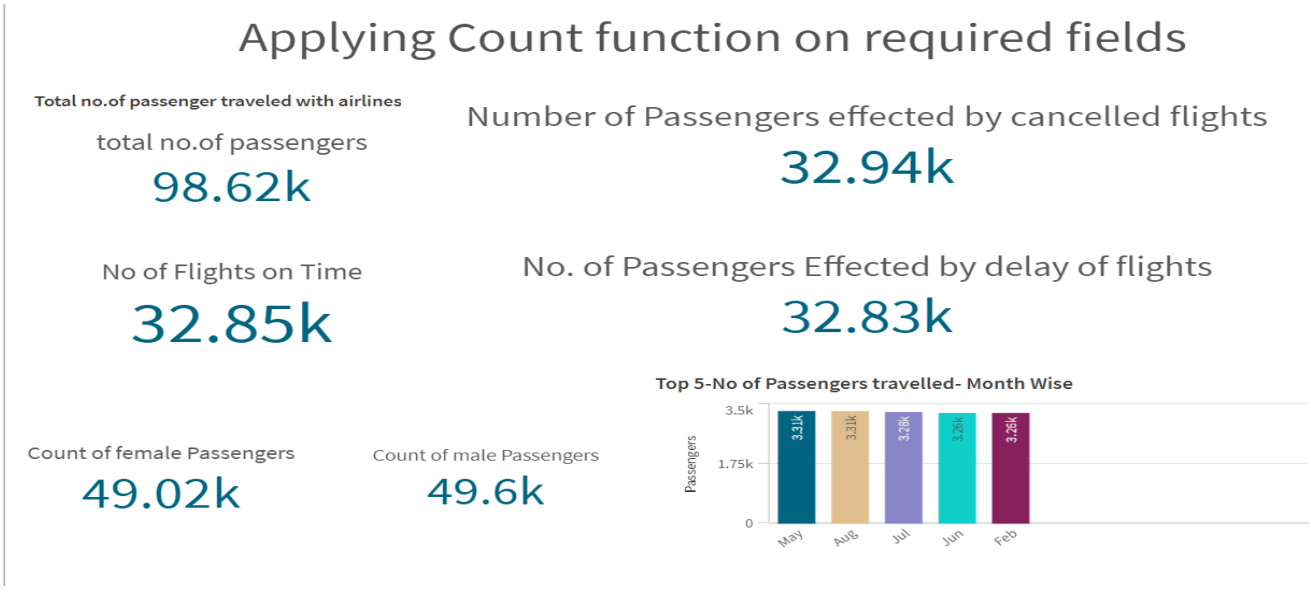
The above Dashboard gives inforamtion about the country which contains more no.of airports and locating them in world map and finding the count of distinct airport name.

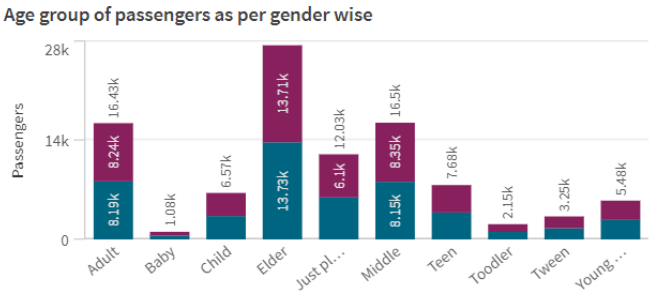


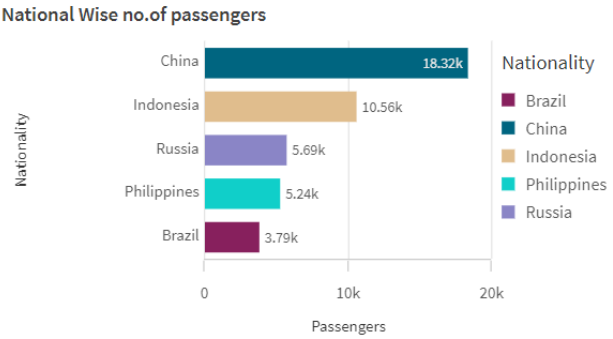
The above Dashboard informs count of flight status based on the filter plane flight status(cancell, delay& On time)and also visualizing count of cancelled , delayed & On time flights for each country.

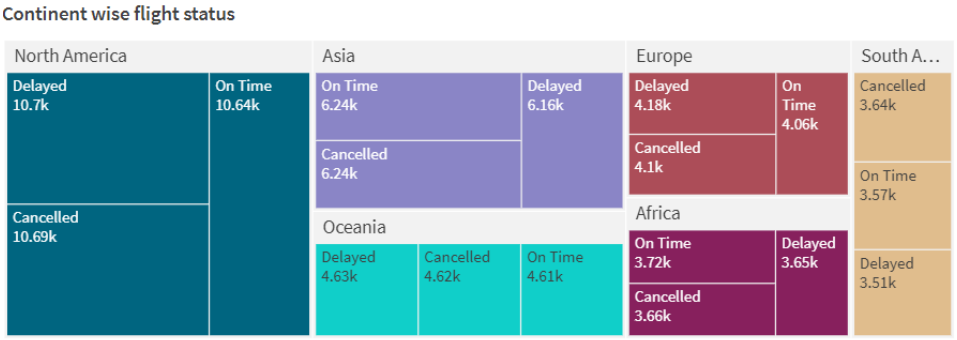
7.REPORT

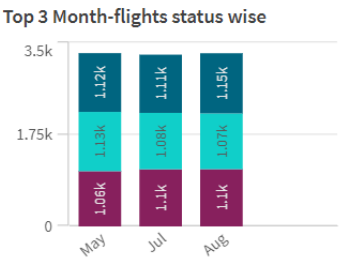
7.1 REPORT CREATION

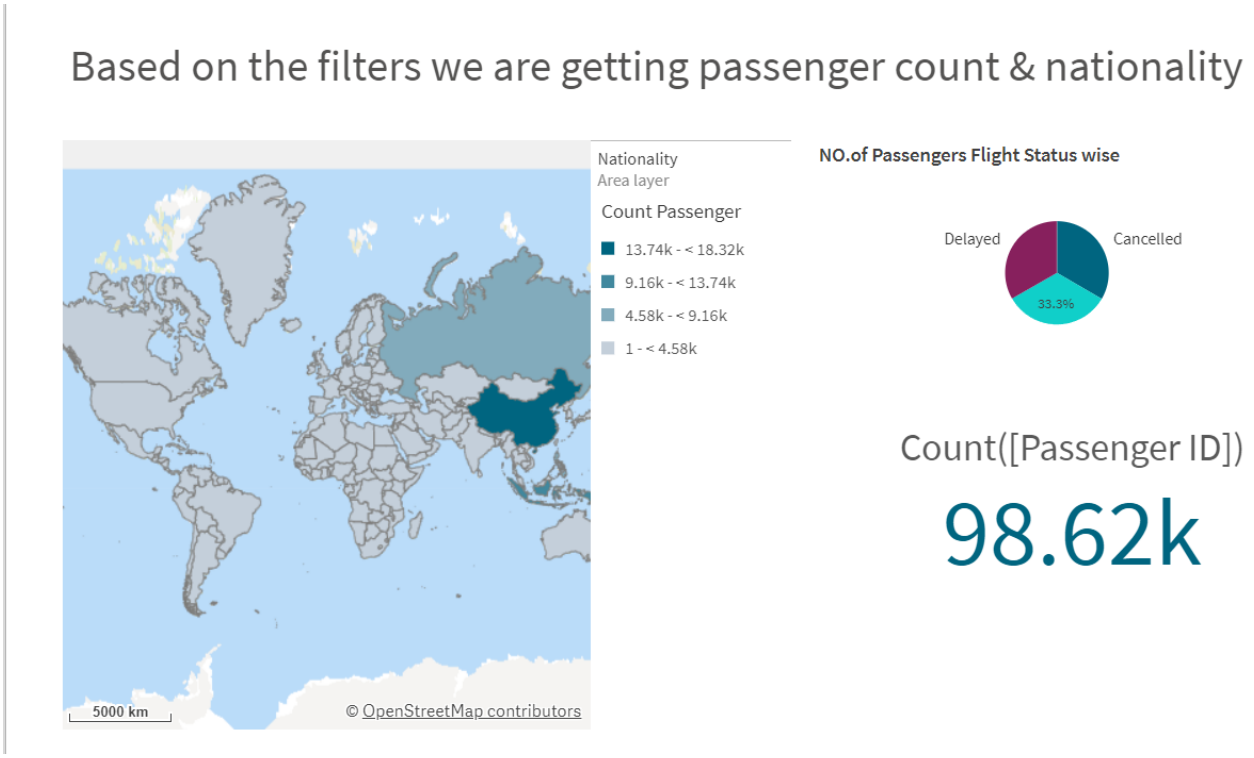






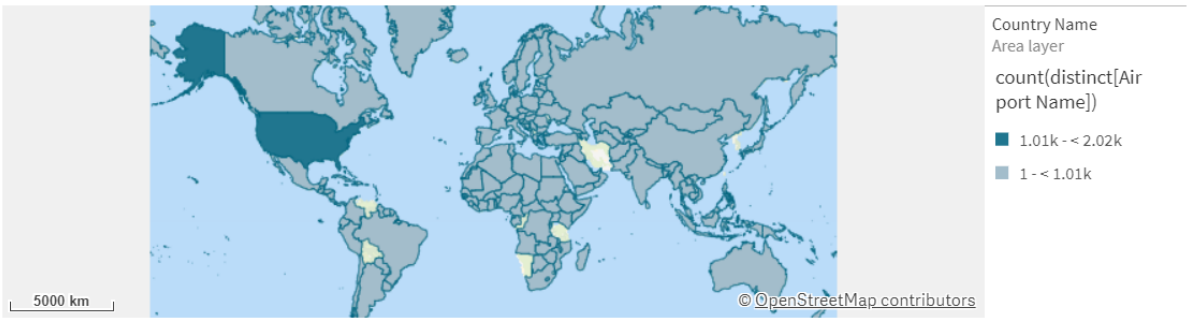






The above figure or Dashboard will changes Based on the applied filters the map and pie chart will be changing. in nationality we getting count of passengers based on groping nationality and pie chart is showing the no.of passengers flight status wise.

finding count distinct airport names and country that have more airports



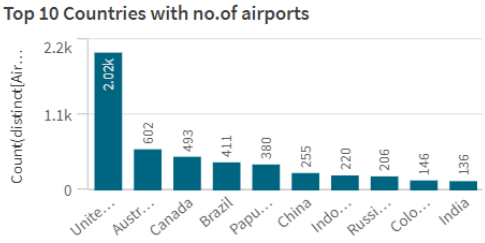
Count of flight status based on the flight status like cancelled, delayed &On time and we count based on the filter we applied

Count([Flight Status])

98.62k

Count(distinct[Airport Name])

9.06k



The above Dashboard gives inforamtion about the country which contains more no.of airports and locating them in world map and finding the count of distinct airport name.

8. PERFORMENCE TESTING

8.1 AMOUNT OF DATA RENDERED

Amount of data used for analysis is 98620 rows and 16 columns

8.2 UTILIZATION OF DATA FILTERS

filter that are used to anaylize the data are

1.Flight Status: this fliter is used now how many passengers are effected by cancelled and delayed flights.

2.Country name: this fliter is used now how many airports are there in that country & also used make count on passengers travelling from perticular country.

3.Airport name: this fliter is used now how many flights reached on time , cancelled & delayed to perticular airport.

4.month: month filter used to now on which month people are travelling more and from which country