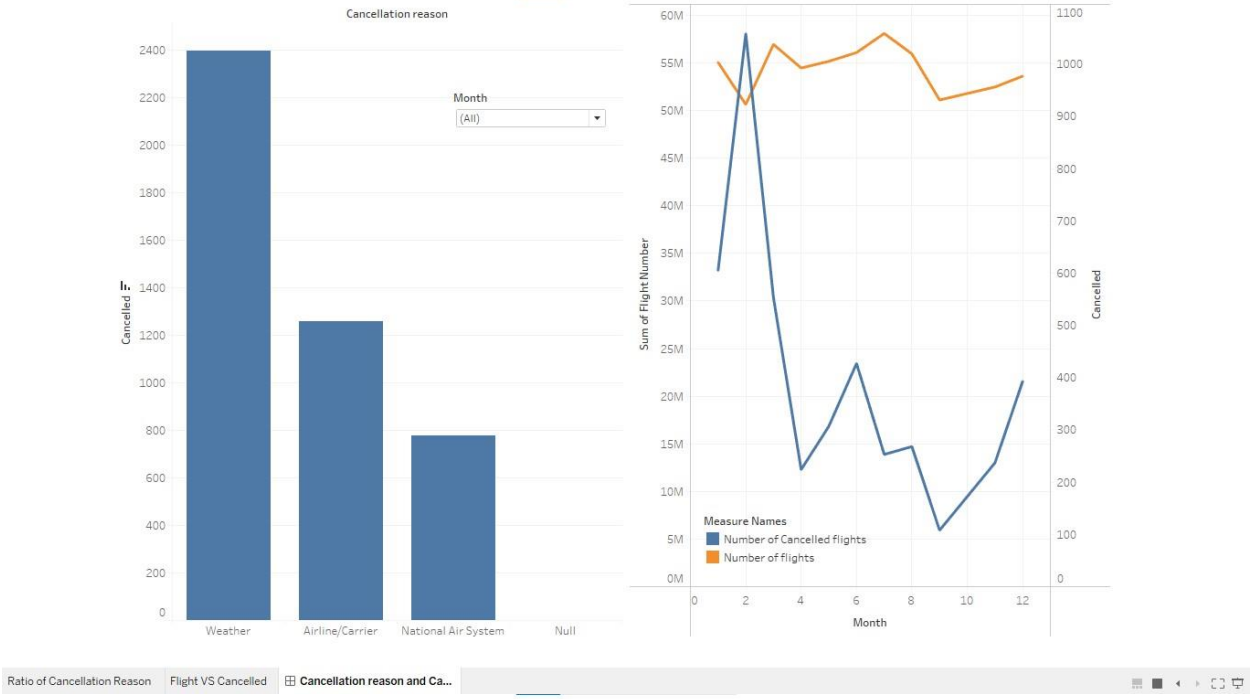
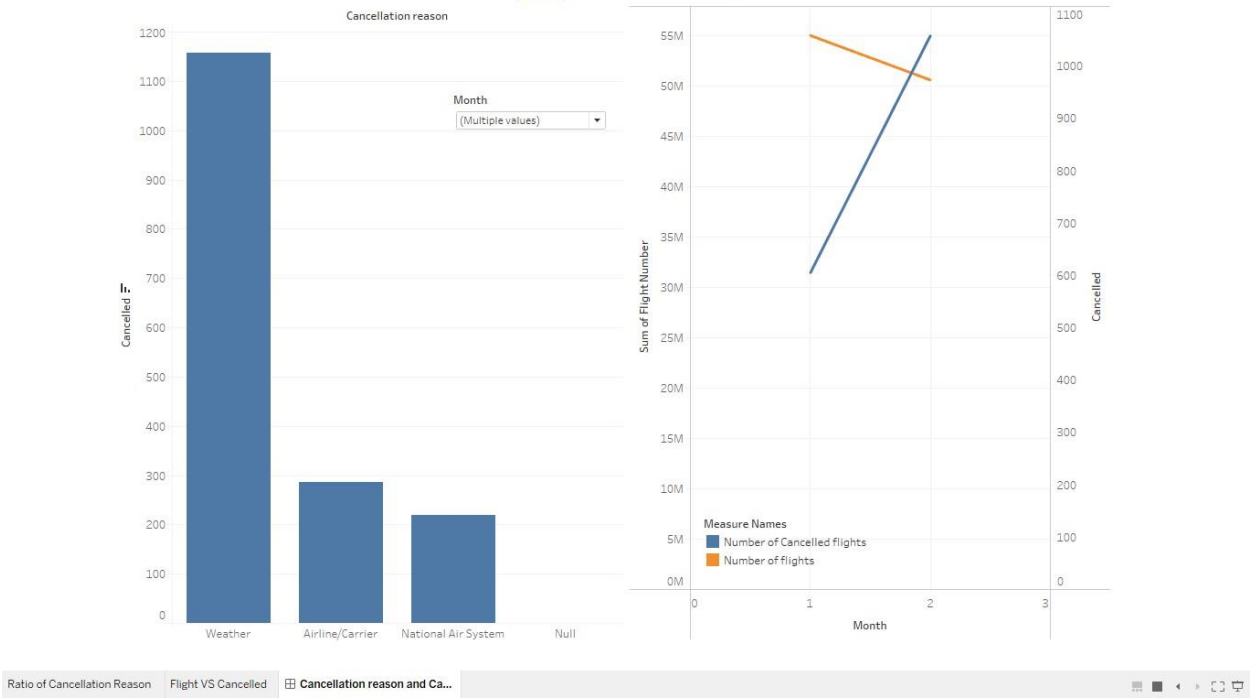
Insight1: Find the reason for the cancellation





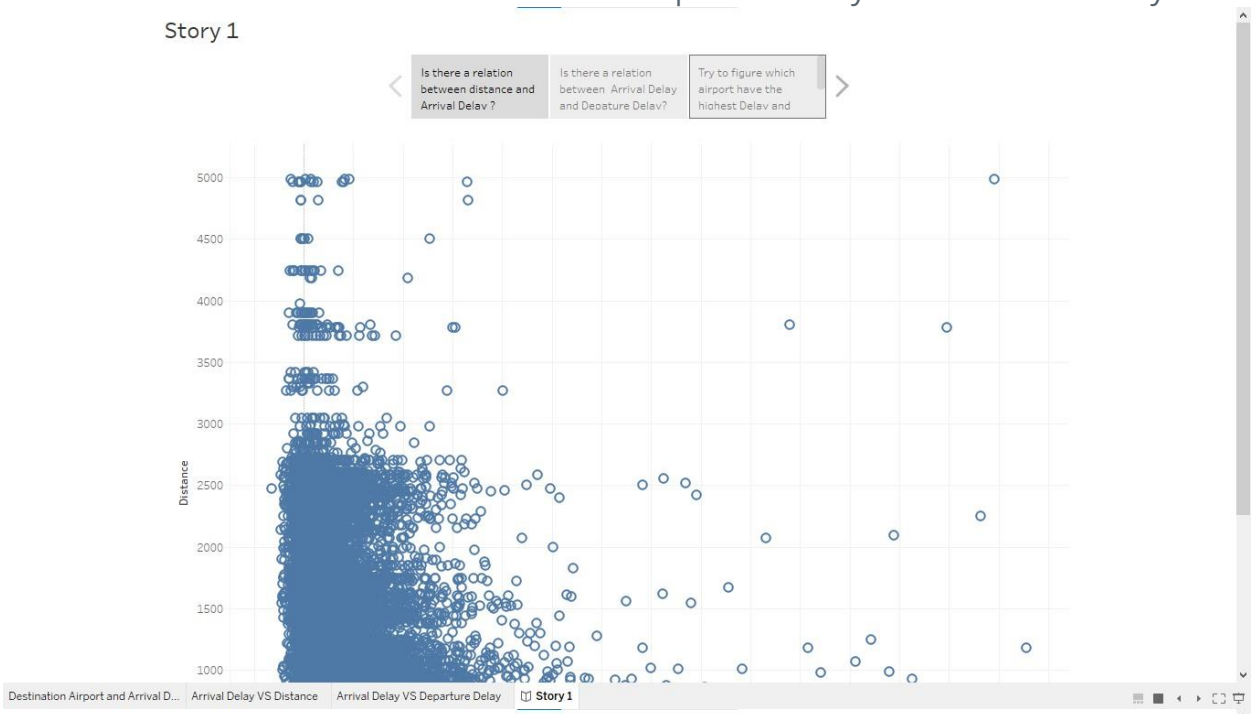
Summary:

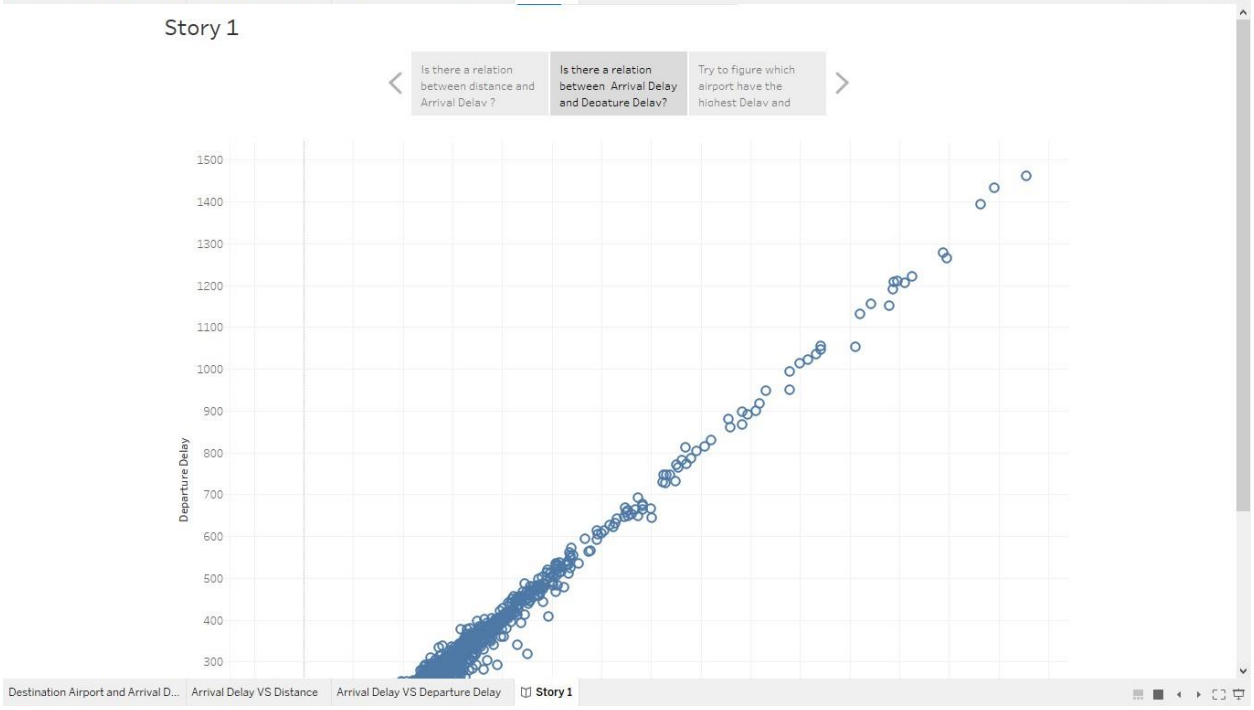
1- In the first sheet we could see that the main reason for the cancellation is the weather 2397 cancelled flights.
2- We try to find the relation between the weather as a reason and the months and we find that February has the largest cancellation flights 1058, and that is clear since February are the coldest month in states.
3- We could notice that event February has the largest cancellation rate it has the least flights over the whole months (50630349).
4- By using the filter, we could see that in other months the reason will change, for example In July the main reason will be airline (133 flights) and the least reason will be weather (54) and that is clear that the weather is much better in July.

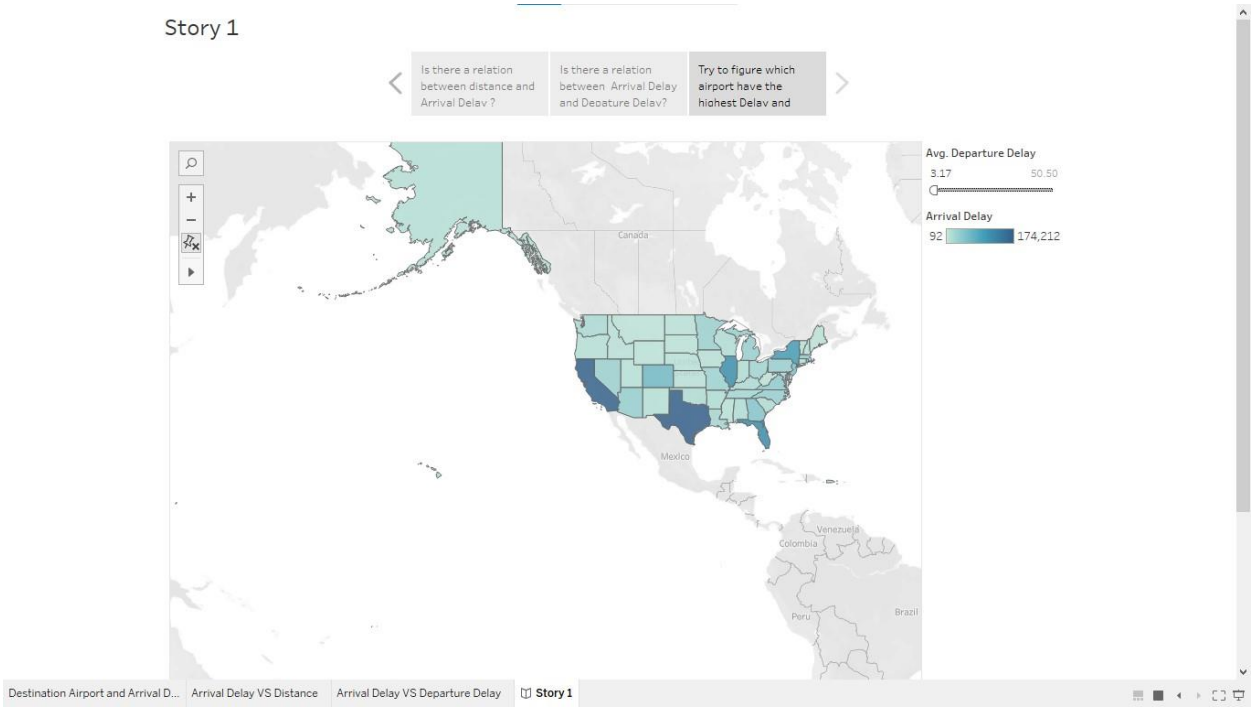
Design:

- first, we replace the values in the Cancellation Reason column by create calculated field (Cancellation reason) and use the readable values instead of coded values.
-we use bar chart for the reason with the number of flights because we deal with categorical data.
- we use line chart to compare between number of flights and number of cancelled flights over months.
-we use filter to compare through months.

Insight2: find the relation between the destination airport and the arrival delay and if there is a relation between distance and the departure delay and the arrival delay.







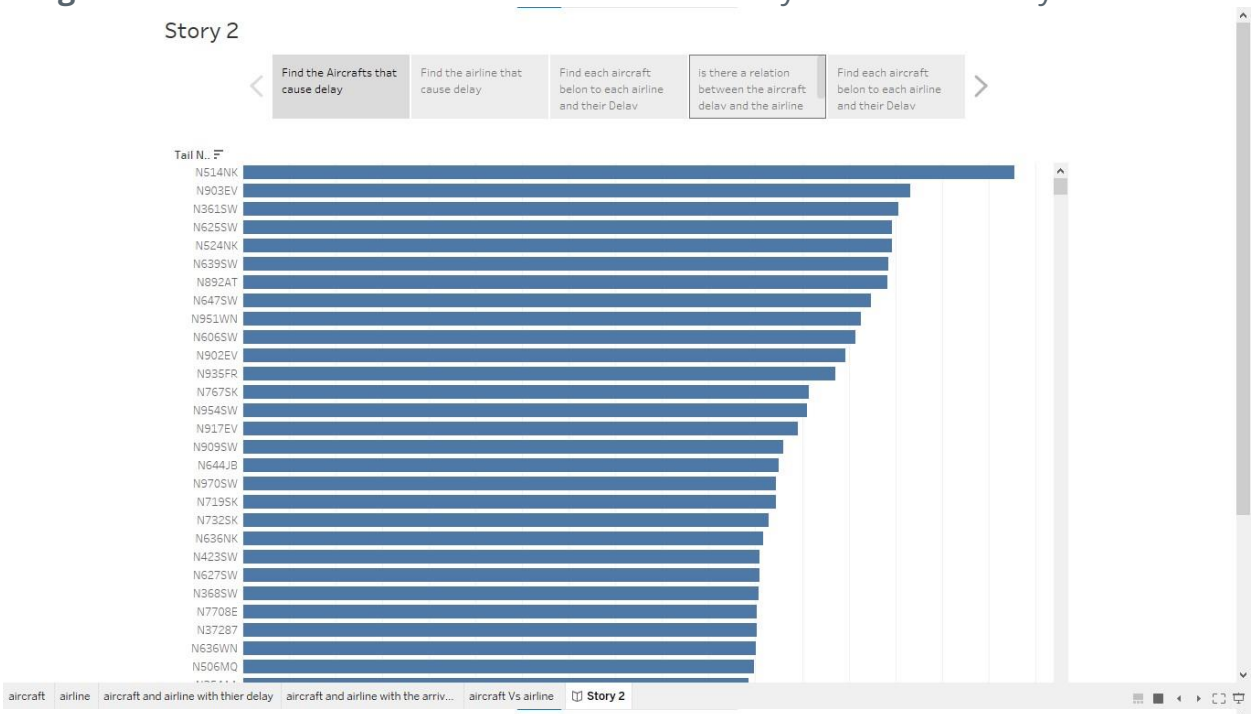
Summary:

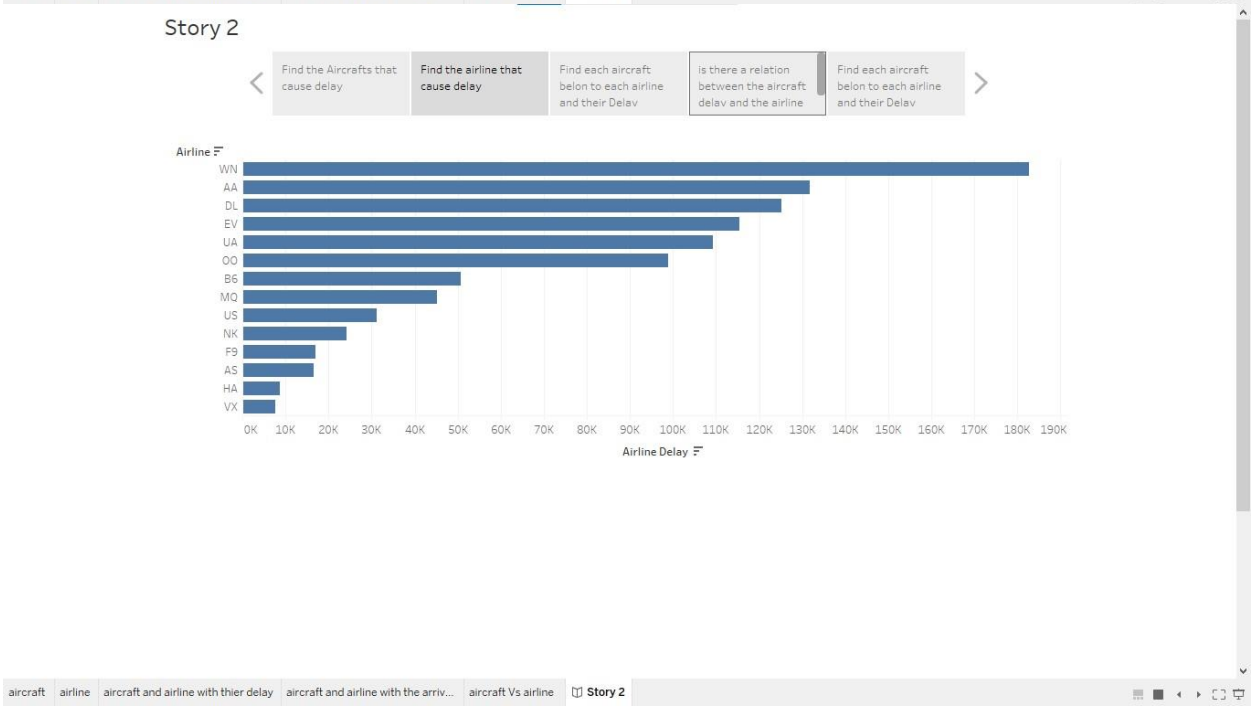
1- We see that Texas has the largest arrival delay (174212) then California (171586).
2- We try to find if there is a relation between distance and arrival delay, but it sounds there is not.
3- We try to find if there is a relation between departure delay and arrival delay, and it sound there is a strong positive correlation between them both.
4- Since there is strong relation between arrival and departure delay, we try to use filter to see if the change of the departure delay will cause a difference according to the states but there is not big different.

Design:

-first, we make an inner join between flights table and airports table on Destination Airport in flight table and Iata code in airports table.
-we use map to trace states.
-we use scatter plots to find the correlation between columns.

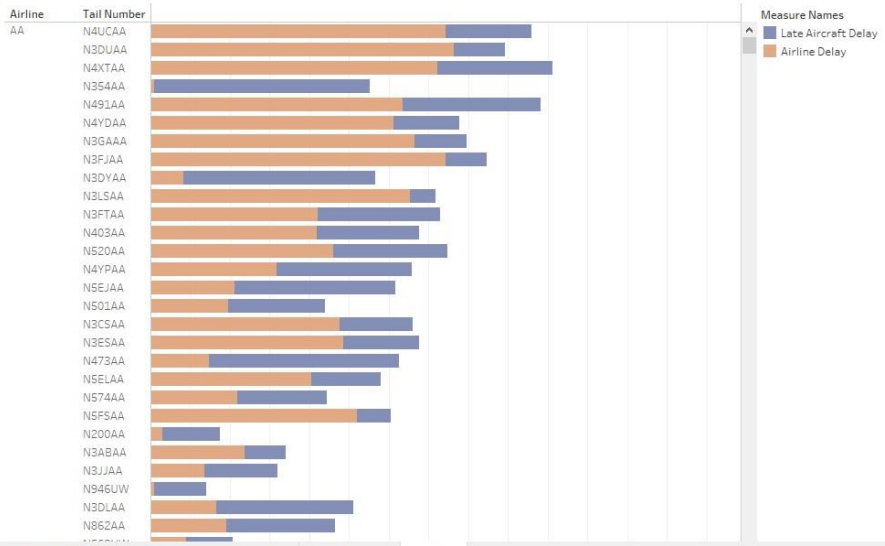
Insight3: find if there is a relation between airline delay and aircraft delay.





Story 2

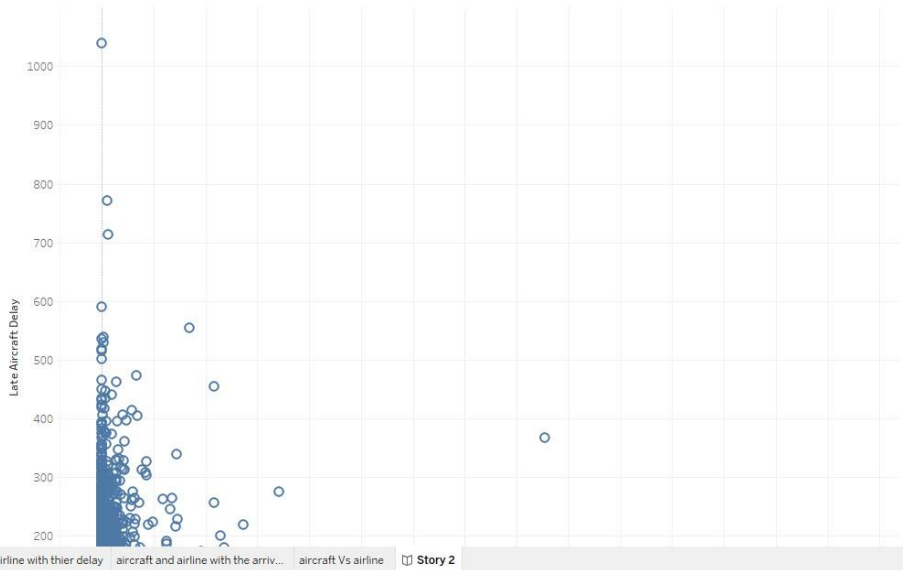
Find the Aircrafts that cause delay | Find the airline that cause delay | **Find each aircraft belon to each airline and their Delav** | is there a relation between the aircraft delav and the airline | Find each aircraft belon to each airline and their Delav



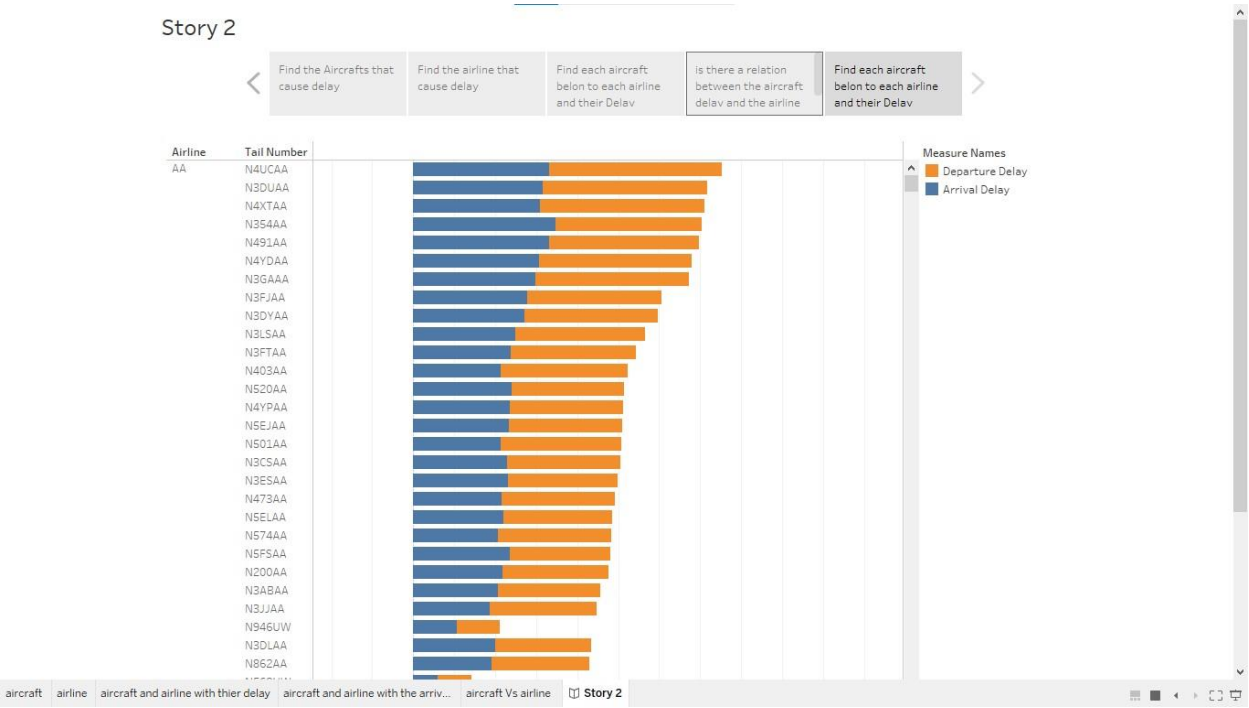
aircraft | airline | aircraft and airline with thier delay | aircraft and airline with the arriv... | aircraft Vs airline | Story 2

Story 2

Find the Aircrafts that cause delay | Find the airline that cause delay | Find each aircraft belon to each airline and their Delav | **is there a relation between the aircraft delav and the airline** | Find each aircraft belon to each airline and their Delav



aircraft | airline | aircraft and airline with thier delay | aircraft and airline with the arriv... | aircraft Vs airline | Story 2



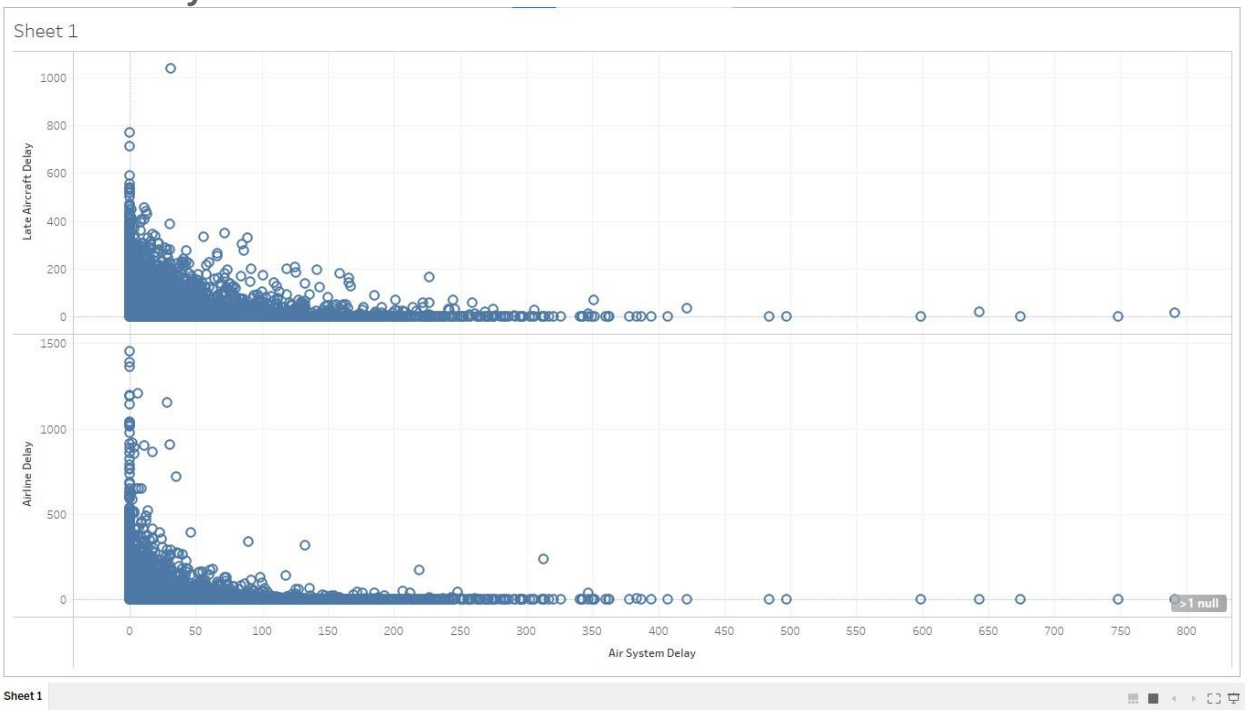
Summary:

1- We try to find the aircrafts that cause the delay (N514NK has the highest aircraft delay with 1655).
2- We try to find airline that cause the delay (WN has the highest airline delay with 182670).
3- We use the hierarchy of airline with its aircraft with the aircraft delay and airline delay.
4- Then we find that there is a week relation between both of them.
5- We try to find whether the airline and aircraft cause the problem in arrival or departure delay, and it sound it cause both of them equally.

Design:

-We create a hierarchy using aircraft and airline.
- we use bar chart since we deal with categorical data.
-we use scatter plot to find the relation.
- we use staked bar since we try to measure two values like arrival delay and departure delay.

Insight4: find if there is a relation between air system delay and airline delay and aircraft delay.



Summary:

We think that may be the problem in the air system could be a problem in a particular airline or aircraft, but we found there is no relation between air system delay and other delays

Design: we use scatter plot to find the relation between them.

Insight5: Find the relation between the day of week and flights and the state visited.

Summary:

1- According to the system in US, Day 1 is Monday and the weekend start form the afternoon of Friday till the afternoon of Sunday.

2- We notice that the week starts with higher number of flights then slightly decrease till Saturday then increase again.
3- We use filter to see if in different days of week that the visited state will affect but it sounds not true and still Texas has the largest number of flights over the week.

Design:

-first, we make an inner join between flights table and airports table on Destination Airport in flight table and Iata code in airports table.
-We use map to trace states.
We use line chart to find the relation between the days and the flights.
-We use filter to try different days.