

Process Book

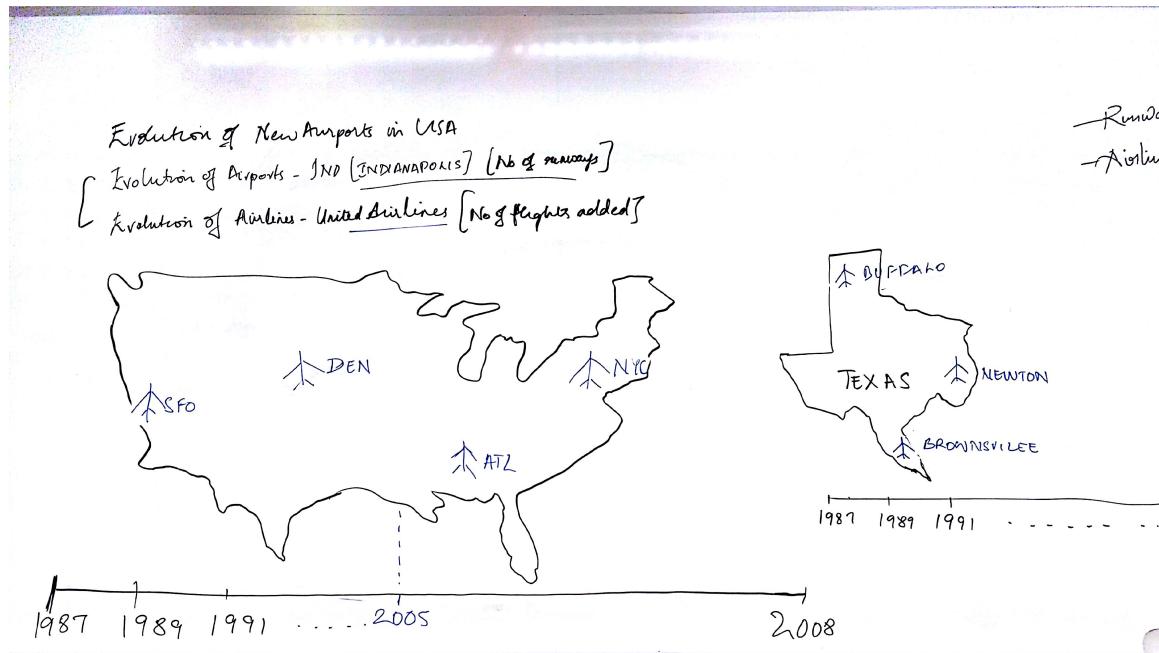
Evolution of Airports across United States of America

- Anusha Ramamurthy
- Santhosh Soundararajan

Initial Ideas

Our initial idea was to create a visualization that would show the onset of airports. We wanted to create small pins over a map of United States, such that each pin represented an airport. Over time it would show how the number of airports increased or decreased, as the map would get more crowded with many pins. After this, clicking on a state would drill down to the spread of airports across that state.

Our initial sketch was as below.



Why a Map?

Maps are a beautiful way to show information about a particular country or location. Since we are displaying information about airports, the best representation was to show flights overlaid on the map.

Picking the dataset

However we couldn't find a dataset that gave us information dating back to the first airport in US. We found a dataset at <http://stat-computing.org/dataexpo/2009/> That had information about flights across different source and destinations, between the years 1987 and 2008.

Hence we decided to change the visualization to show airports across US and how they got busy with time.

Our first real challenge was processing the dataset. We had close to 12 GB of data in csv files. We decided to use python for processing the data.

We first tried to find out how many unique airports were there in our dataset. We realized that for the year 1987 alone, we had 230 unique airports!

```
In [2]: import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import seaborn as sb
%matplotlib inline

In [3]: df = pd.read_csv('1987.csv', delimiter=',', engine='python')

In [4]: origins = df['Origin'].unique()

In [5]: len(origins)
Out[5]: 237
```

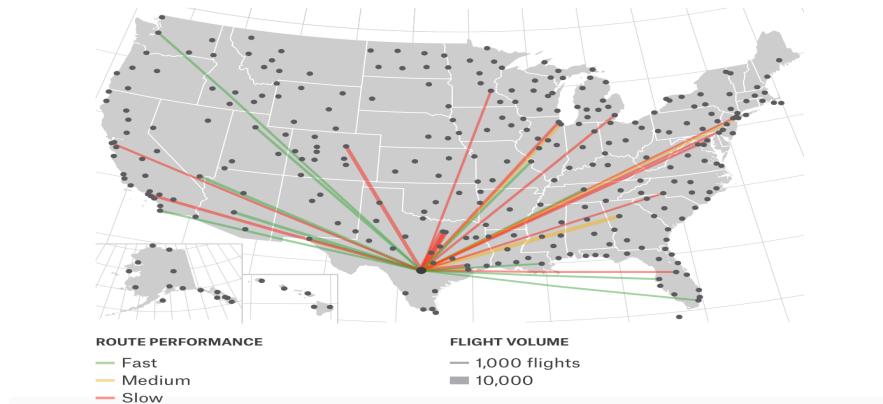
Plotting 237 airports on a map would make the map look extremely cluttered so we decided to show top 15 busiest airports.

We had to get the top 20 airports, the total incoming and outgoing flights for each year and map it to a latitude and longitude location such that we could visualize them on the map. Since the data was over 12 Gb, the processing took close to 3 hours. Writing the code and the methods to process data took us quite long. But python and pandas library made handling data easy.

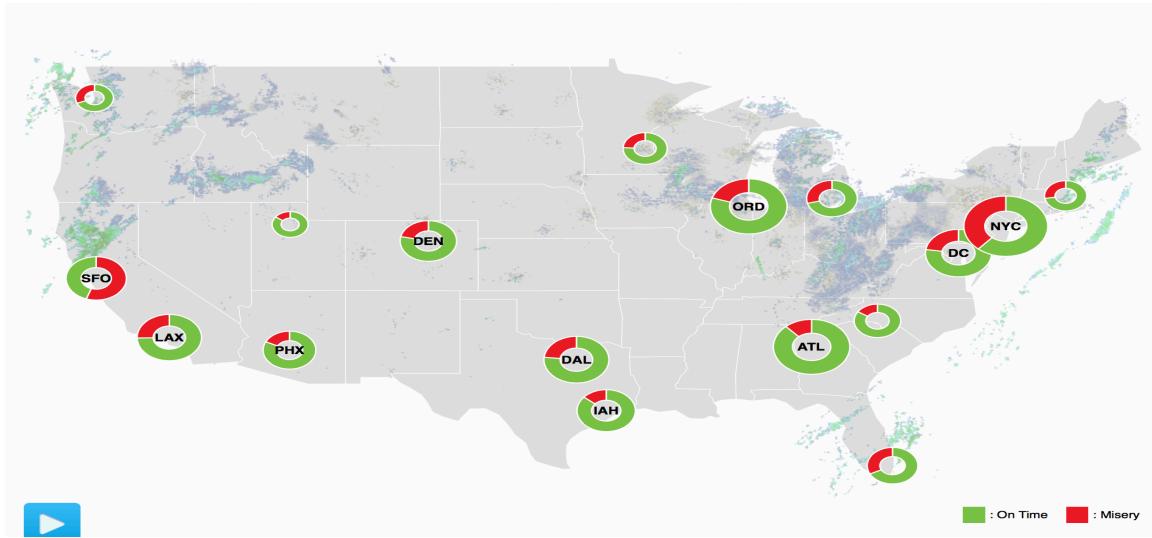
Visualization:

We looked at some of the related visualizations such as:

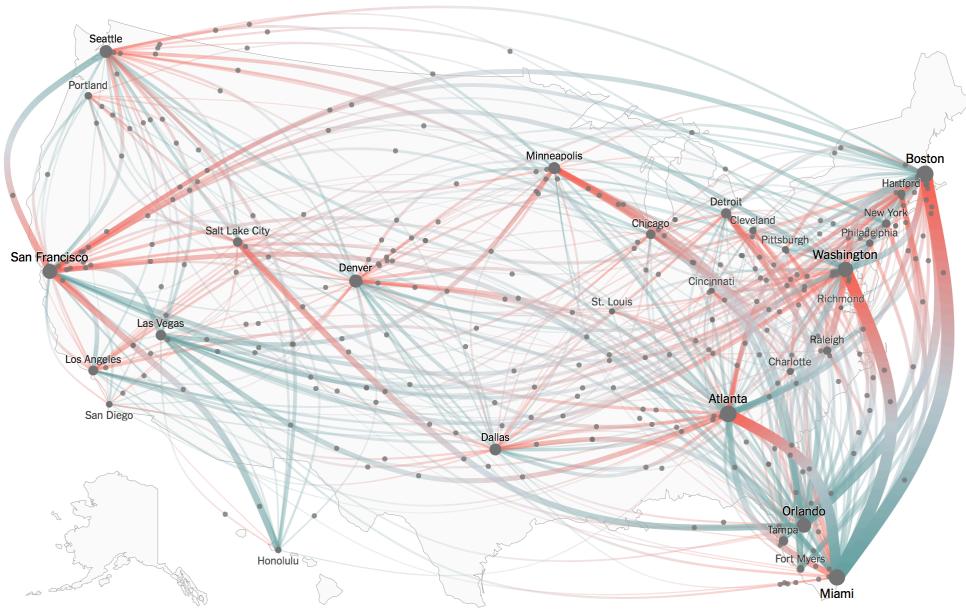
- <http://projects.fivethirtyeight.com/flights/>



- <http://flightaware.com/miserymap/LIMG/1450474200>



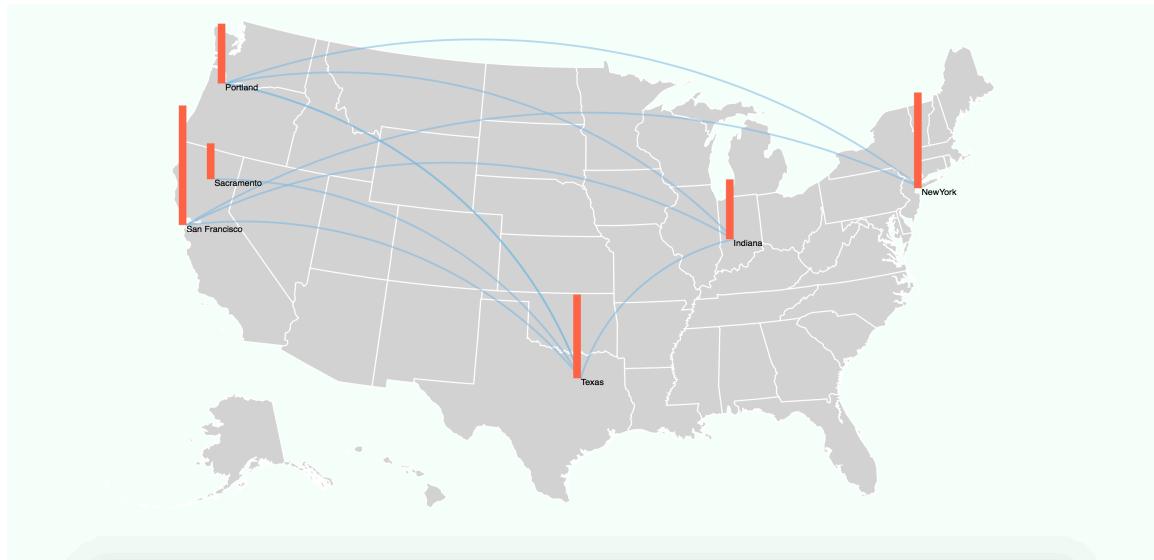
- http://www.nytimes.com/interactive/2015/11/24/upshot/thanksgiving-flight-patterns.html?rref=upshot&smid=tw-upshotnyt&smtyp=cur&_r=1



However all these visualizations show the path taken by different flights. Since we wanted to depict the increase in traffic at an airport, we had to find a better way to visualize this.

After discussions, we decided to put bars to show the increase in the magnitude of flights over time.

Our initial work looked as below.



The paths connecting the locations were not required, since they were confusing the message we wished to convey.

We wanted a way to show the value of incoming and outgoing flights at an airport and how it has changed over the years. On discussion with YY, he suggested we look at this visualization by NY times, which puts across trend information in very succinct manner.

<http://www.nytimes.com/interactive/2014/06/05/upshot/how-the-recession-reshaped-the-economy-in-255-charts.html>

How the Recession Reshaped the Economy, in 255 Charts

By JEREMY ASHKENAS and ALICIA PARLAPIANO Updated: JUNE 6, 2014



We added a tool tip at top 20 airports to show the total incoming and outgoing flights, with trend chart to show how the value has changed over the years.

We had problems integrating all the layers. Since we had created one svg layer for each component, like one for the maps, one for the bars and one for the tool tips. We felt that the a timer function, like a count down for the 21 years with flight counts increasing would also help the viewer get a better idea.

Final visualization



Future Work

We would like to create more interactive elements. We also like to show the total delay at each airport and how this value has improved or worsened at different airports.