



# United States Flight Aggregation

Trevor Grayson and Ted Gundlach



# Topic and Introduction

Air traffic, and transportation in general, is one of the most important facets of human society and requires heavier management than most of society's "systems".

Even with this importance, many people never stop and consider how it all connects. So, because of this importance, this seemed like a cool way to explore big data.

- Periodically track all flights across the United States
- Determine the most heavily traveled "air lanes"
- Determine the most heavily used airports
- View other variables available such as trajectory velocity



# Methods

## Periodic Data Collection

- Scheduler and timer modules
- Collection of flight data every minute for 2 days via OpenSky API
- Stored in database later translated into parquet file to be called by DataShader

## DataShader

- Trajectory pathing (ascending, descending)
- Velocity pathing (airspeed across trajectory)

# Data Collection

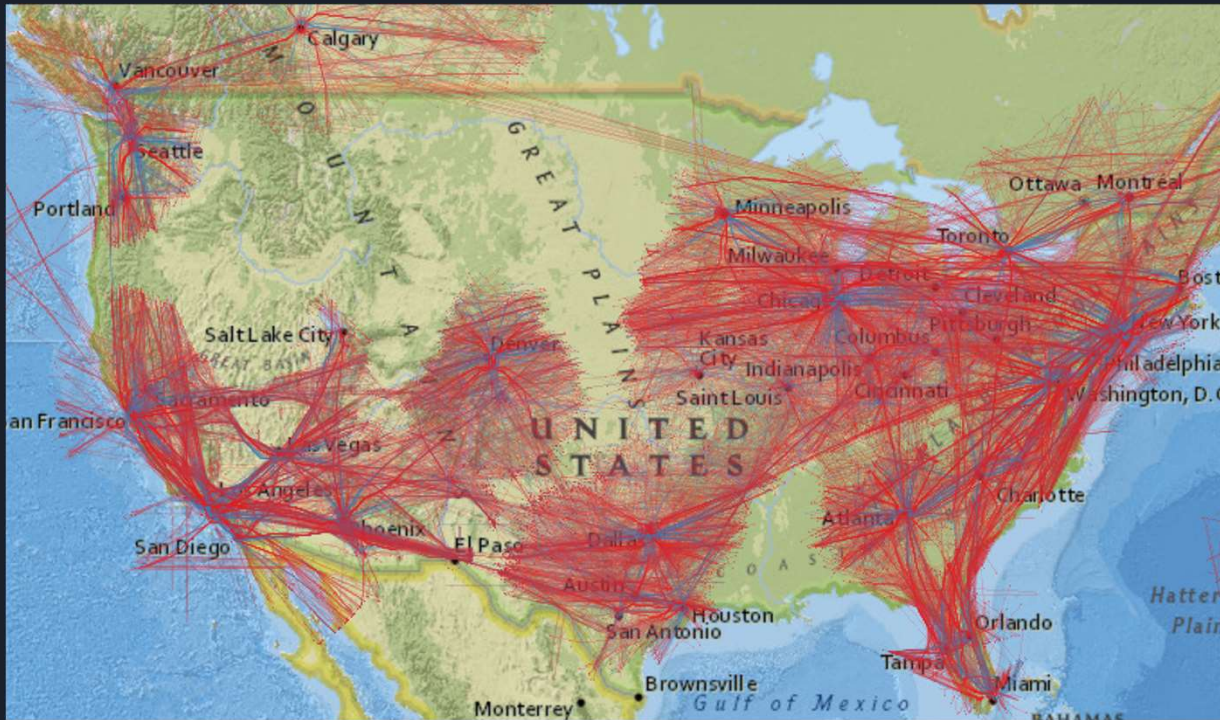
Tables (1)

opensky

```
CREATE TABLE "opensky" ( "icao24" TEXT, "callsign" TEXT, "origin" TEXT, "time_position" REAL, "time_velocity" I
```

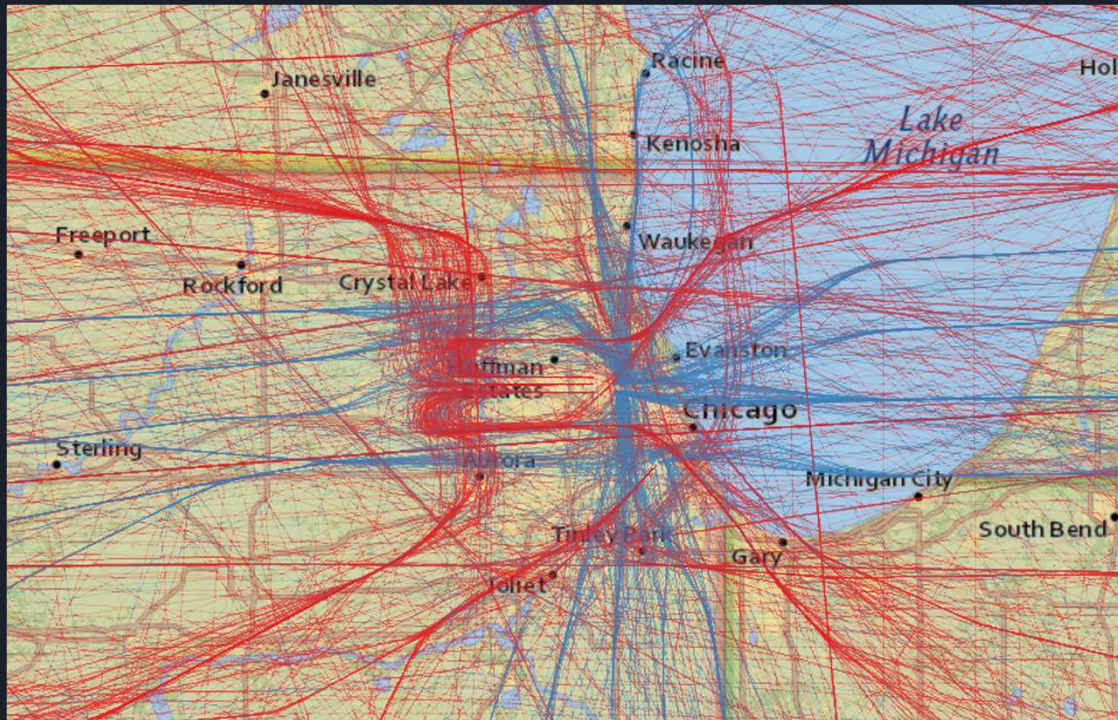
icao24	TEXT	"icao24" TEXT
callsign	TEXT	"callsign" TEXT
origin	TEXT	"origin" TEXT
time_position	REAL	"time_position" REAL
time_velocity	INTEGER	"time_velocity" INTEGER
longitude	REAL	"longitude" REAL
latitude	REAL	"latitude" REAL
altitude	REAL	"altitude" REAL
on_ground	INTEGER	"on_ground" INTEGER
velocity	REAL	"velocity" REAL
heading	REAL	"heading" REAL
vertical_rate	REAL	"vertical_rate" REAL
sensors	TEXT	"sensors" TEXT
misc1	REAL	"misc1" REAL
misc2	TEXT	"misc2" TEXT
misc3	INTEGER	"misc3" INTEGER
misc4	INTEGER	"misc4" INTEGER
timestamp	INTEGER	"timestamp" INTEGER

# Ascending and Descending (United States)

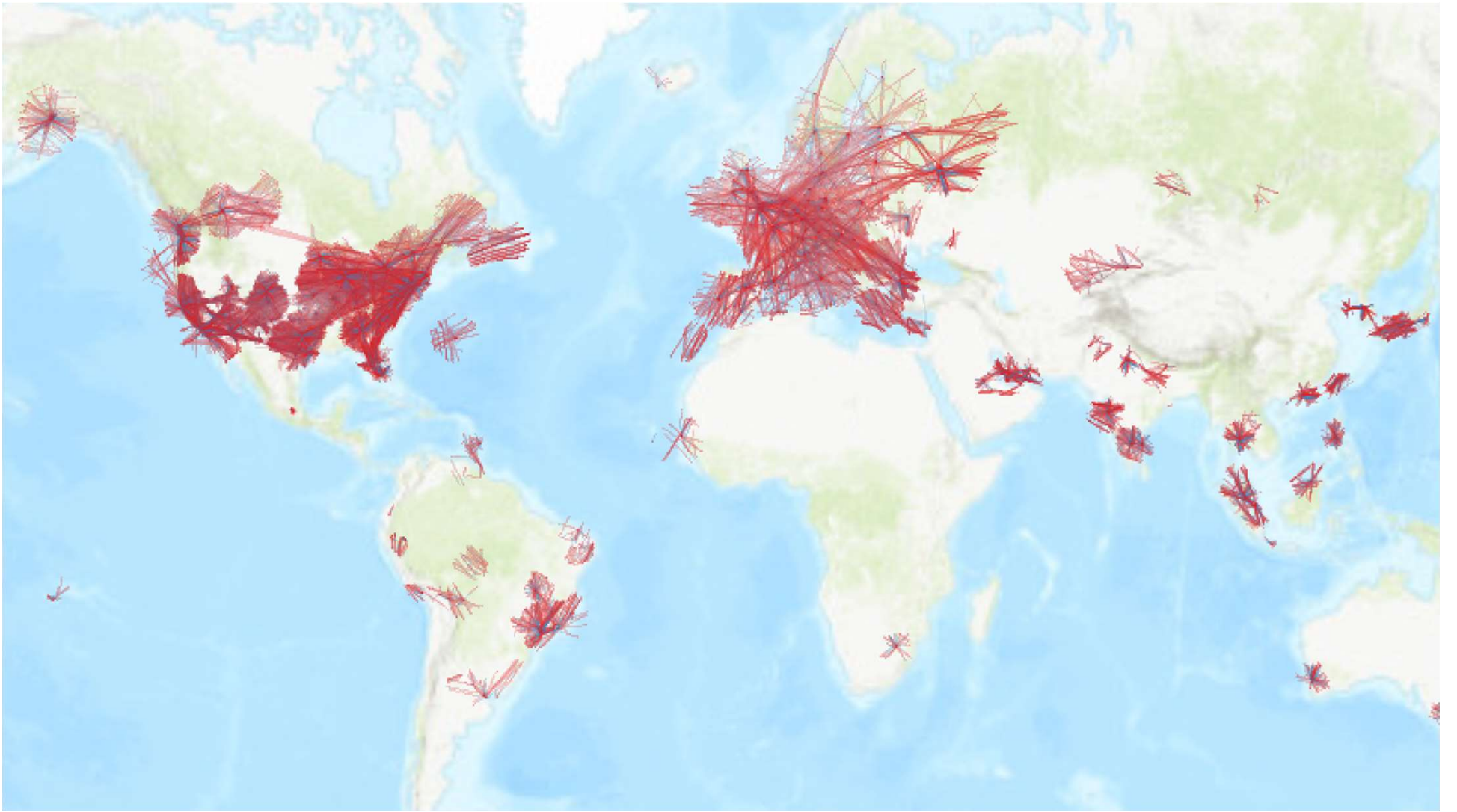


- Red = Descending
- Blue = Ascending
- Blank Spaces?
- Clusters

## Ascending and Descending (Chicago - O'Hare, Midway)



- Red = Descending
- Blue = Ascending
- Takeoff and Landing Patterns





# Velocity Pathing

Goal was to obtain velocity along trajectories to see patterns of slowing to descend and acceleration at ascension.

However, we do not own strong enough computers to do the necessary computations needed for our data size.

299

@staticmethod

**MemoryError:**





## Memory Issues

Only able to use a half a days of data, since our 2 days file was too large to convert for the memory we had available to us on our laptops.

Just for a reference, Open-Sky has been up for just over 2,000 days and has collected over 14 trillion pieces of data from aircraft. They receive over 300,000 pieces of data every second.

As with velocity, majority of variables required too much computational power to be visualized. This made it difficult to spot any other trends that may have come from data.



## Conclusions & Discussion

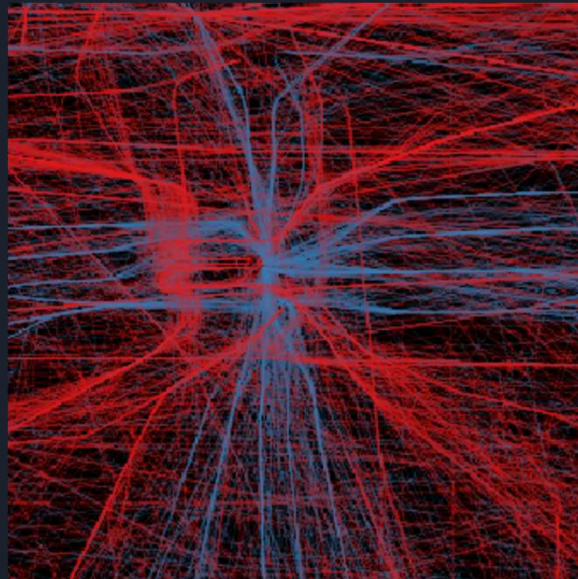
- Descending distance > ascending distance
- Utilization of specific runways to land whereas taking off is related to direction of destination
- Largest airports are easily noticed (ORD, JFK, LAX, SFO, etc.)
- Highest concentrated flight lanes lie on the east coast (most traffic)
- United States has more air traffic than any other country or area on the planet
- **Big Data** consumes a lot of space and requires a lot of computing power
- A lot more air traffic than expected
- Blank sectors (no climb or drop in altitude?)

# Acknowledgments

Dr. Song Gao

DataShader (<http://datashader.org/>)

OpenSky-Network (<https://opensky-network.org/>)



Questions?