**Objective-** Should we start a ride hailing (taxi/limo) /ride sharing service in San Diego, CA or Dallas, TX? The recommendation is based in terms of demand, competition, expected wage and safety.

Taxi driving in NYC- Taxi driving is well defined and regulated in cities like New York and Chicago. We collected yellow taxi data to understand recent trends in fares and miles driven. See fig below-

A screenshot of a cell phone

Description automatically generated

**Insight:**

*On an average, a taxi driver drives 3 miles and charges around $15 per trip*

Source: <https://opendata.cityofnewyork.us/>

In order to start a ride haling/ride sharing service in San Diego or Dallas, we investigated the following-

1. **Demand-**

In our analysis, we consider population size as a proxy for demand and their median ages as the typical demographic attribute of taxi/rideshare (uber/lyft) users.

According to Statista (Link- <https://www.statista.com/statistics/822833/us-ride-sharing-uber-users-age/>) and global web index (Link- <https://blog.globalwebindex.com/chart-of-the-day/uber-demographics/>), 16 to 34 years are the typical demographic attribute of uber/ride share users.

See Table 1 & 2 for county wise population and their comparison to state population.

A screenshot of a cell phone

Description automatically generatedTable1: Below table suggests San Diego & Dallas counties have similar state percentage populations.

Table2: Below table suggests similar median ages of current population in both the counties

A screenshot of a cell phone

Description automatically generated

**Insight:**

*The median ages of San Diego and Dallas counties, as well as their current population size suggests us to expect “Good” demand in both counties.*

Source: <https://suburbanstats.org>

Procedure: We considered county wise and state wise data to build a general DataFrame. A percentage of State column was generated for comparison purpose.

1. **Competition & Expected Wage-**

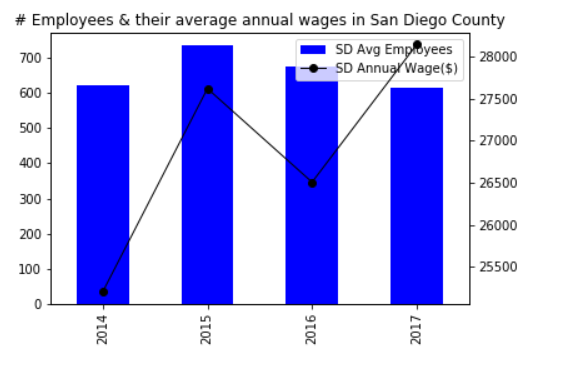
We considered labor statistics data to understand the historical trend in average number of establishments and average number of employees in the taxi and limousine industry.   
See Fig 1: Taxi and Limousine- Avg # Establishments.

A screenshot of a cell phone

Description automatically generated

We also considered the historical average annual wages with # employees. See figures below.

A screenshot of a cell phone

Description automatically generated

**Insight:**

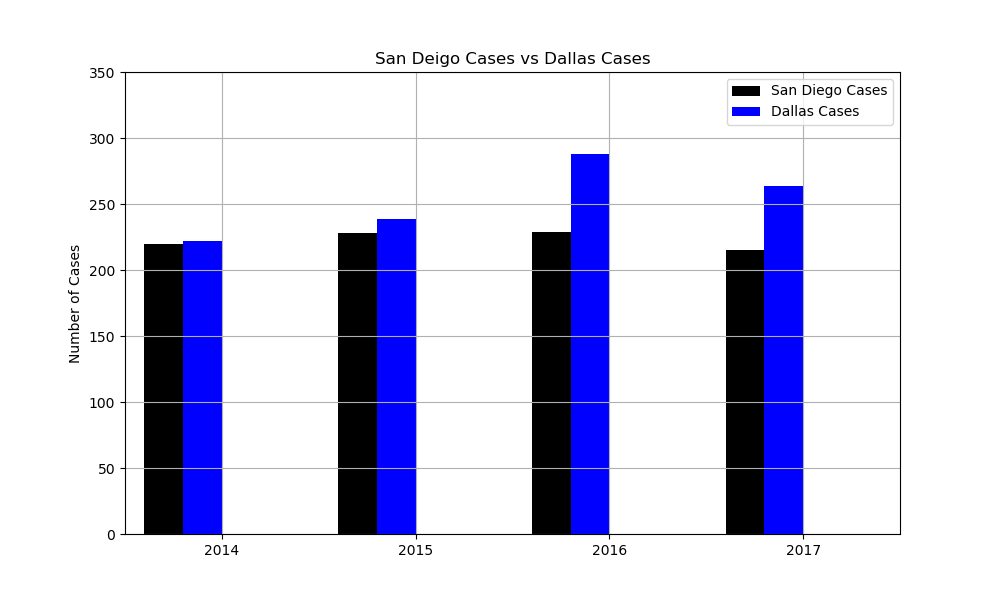
*Above figures suggests less competition with better wages in Dallas over San Diego.*

Source: <https://www.bls.gov>

Procedure: county wise business data was downloaded in CSV format and we formed a DataFrame for plotting graphs.

1. **Safety-**

We considered crash data from NHTSA by performing an API call to understand the # cases registered and fatalities in each county. See below figures.



A screenshot of a cell phone

Description automatically generated

**Insight:**

*Looks like San Diego is safer place for driving as compared to Dallas.*

Source: <https://crashviewer.nhtsa.dot.gov>

Procedure: API call to get crash-data at a county level and then plot graphs using matplotlib

**Final Recommendation-**

I’d consider starting a “safe” ride hailing/ride share service in Dallas county on account of less competition and higher wages.