Project Idea

I. Project Title
Las Vegas: The Strip or Downtown? – Comparative Analysis of Restaurants in Old and New Las Vegas

II. Team members
1. Veronica Lee
   - veronica.lee@columbia.edu
2. Eileen (Yei Rim) Suh
   - ys2595@columbia.edu

III. Abstract
There are two main districts that travelers to Las Vegas visit: the Downtown or Strip Las Vegas. The Downtown Las Vegas (DTLV) is also often called as Old Las Vegas. As the later name implies, it is the historic center of Las Vegas, Nevada. All the casinos and hotels are relatively close together, mostly in walking distance. They are also less expensive that those in the Strip Las Vegas as it requires less minimum bet for different table games. For this reason, DTLV is "for the budget minded and the serious gambler who do not like all the glitz [and] glamour of the Strip [Las Vegas]" [1.]

The Las Vegas Strip, on the other hand, features mega resorts that one often sees in television shows or movies. It has "Disney-like attractions and nothing less than 5 star hotels"[2] All the famous casino and resorts like the Bellagio, Wynn, Venetian, and Mandalay Bay could be found in the Strip Las Vegas. Thus sightseers or less serious gamblers would prefer to stay on the Strip.

We are interested in comparing restaurants in these two regions of Old and New Las Vegas. We will use the dataset that Yelp provides on different business and service provided within Las Vegas region. With this dataset, we plan to present our data analysis through different sets of data visualization including the following but not limited to:
1. Geospatial analysis
   - Layout all restaurants in Las Vegas that are in the Yelp dataset. We would like to see how restaurants are visually distributed in the map.
   - Conduct comparative analysis on high rated restaurants in Old and New Las Vegas region. We are interested in how high rated restaurants in two different regions differ in terms or their business attributes.
   - Compare and contrast what makes visitors of two regions different. For instance, we are interested in whether people visiting restaurants in Old Las Vegas less concerned about ambience compared to those in New Las Vegas

2. Textual analysis

---

[1] "Las Vegas - Downtown Vs The Strip." Singapore Travel Guide. Christopher Allen, Web. 10 Mar. 2017.
[2] "Las Vegas: The Strip or Downtown?" Reviews of Hotels, Flights and Vacation Rentals. N.p., n.d. Web. 10 Mar. 2017.

- Conduct sentiment analysis of high rated restaurants in Old and New Las Vegas region.
- Find correlation between high review & certain key words of reviews by neighborhood in Old and New Las Vegas
- Make word clouds for each star review (from one to five) in two different regions. We are interested in examining whether two different sets of word clouds would look disparate or similar.

IV. Data
We plan to use the Yelp's academic dataset. This dataset has 4.1 million reviews by 1 million users for 144,000 businesses. It provides dataset for different cities like Edinburgh, U.K; Karlsruhe, Germany; Montreal and Waterloo, Canada; Pittsburgh, Charlotte, Urbana-Champaign, Phoenix, Las Vegas, Madison, Cleveland, USA. The dataset also includes 1.1 million business attributes like hours, parking availability, and ambience.

For our project, we are only interested in the restaurant data in Las Vegas in the United States. Thus we plan to use subset of this large dataset.