

YELP DATASET CHALLENGE

An exploratory visualization of
Round 7 Dataset

Deepak Khurana

YELP

A CROWD BASED LOCAL BUSINESS REVIEW COMPANY

Punjabi Grocery & Deli

Is this your business?



305 reviews

[Details](#)

[★ Write a Review](#)

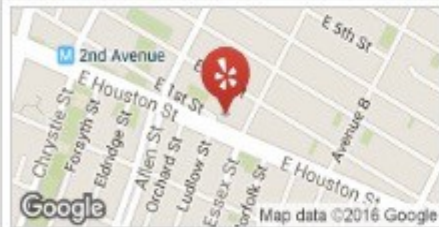
[Add Photo](#)

[Share](#)

[Bookmark](#)

\$ • Indian, Grocery, Vegetarian

[Edit](#)



114 E 1st St
New York, NY 10009
b/t 7th St & 6th St
East Village, Alphabet City

[Edit](#)

[Get Directions](#)

[2 Ave. and 2 more stations](#)

[\(212\) 533-3356](#)



"Their chana masala (**chickpeas**), saag (spinach) and daal (lentils) are always there and always amazing." in 37 reviews



"They do microwave EVERYTHING (in **styrofoam**) so if you have issue with that eat it cold or stop being such a wuss." in 33 reviews



"This is where all the **cabbies** eat ad if that doesn't tell you hot awesome it is I don't know what will." in 16 reviews

[Show more review highlights](#)



Today **Open 24 hours** [Open now](#)



[Full menu](#)



Price range **Under \$10**



Vegetarians go nuts for this spot. See other places they like.



[Work here?](#) [Claim this business](#)

Hours

Mon **Open 24 hours**

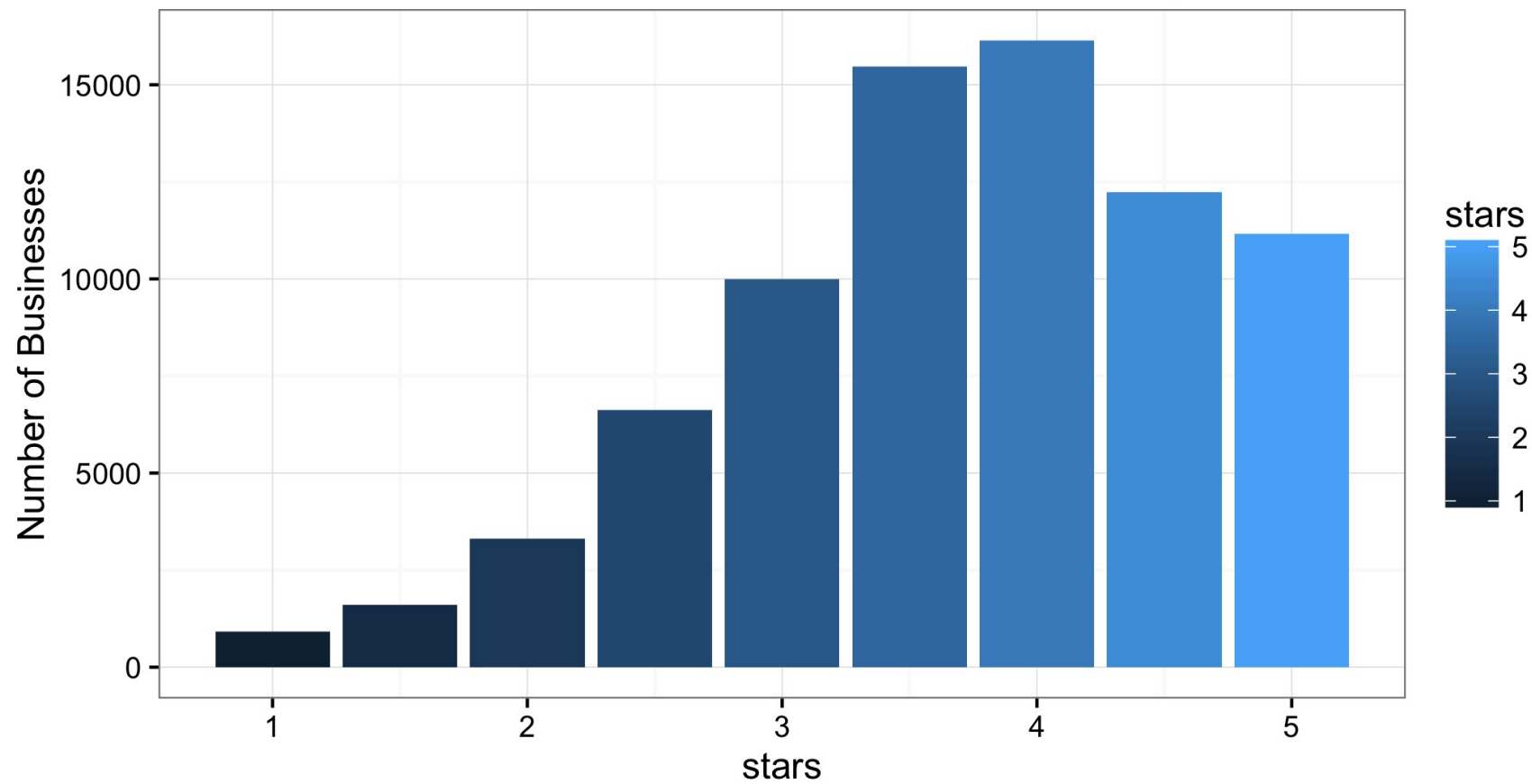
A VERY RICH DATASET FOR ML

The Challenge Dataset:

- **2.2M** reviews and **591K** tips by **552K** users for **77K** businesses
- **566K** business attributes, e.g., hours, parking availability, ambience.
- Social network of **552K** users for a total of **3.5M** social edges.
- Aggregated check-ins over time for each of the **77K** businesses
- **200,000** pictures from the included businesses

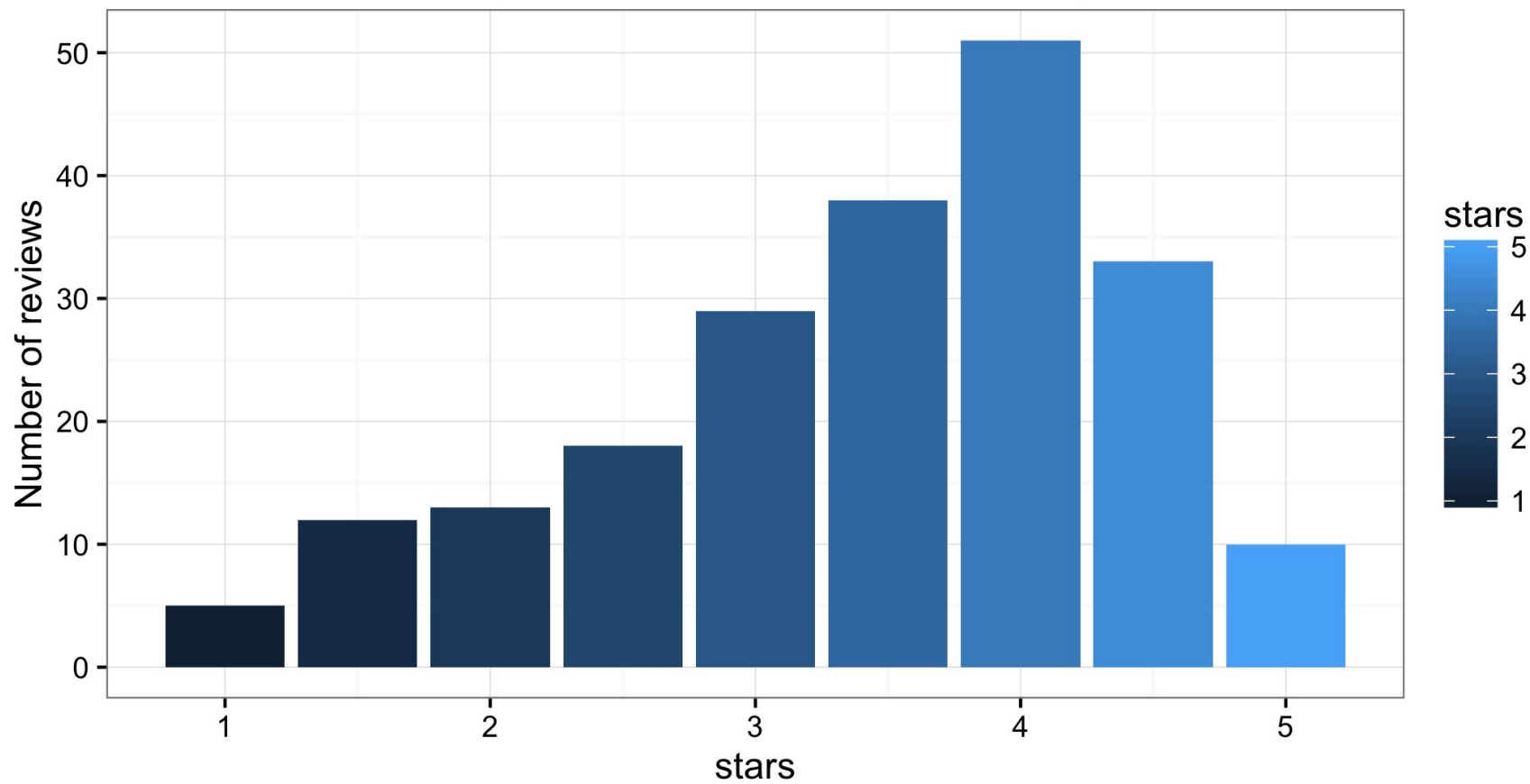
Business Statistics

Business distribution as a function of ratings



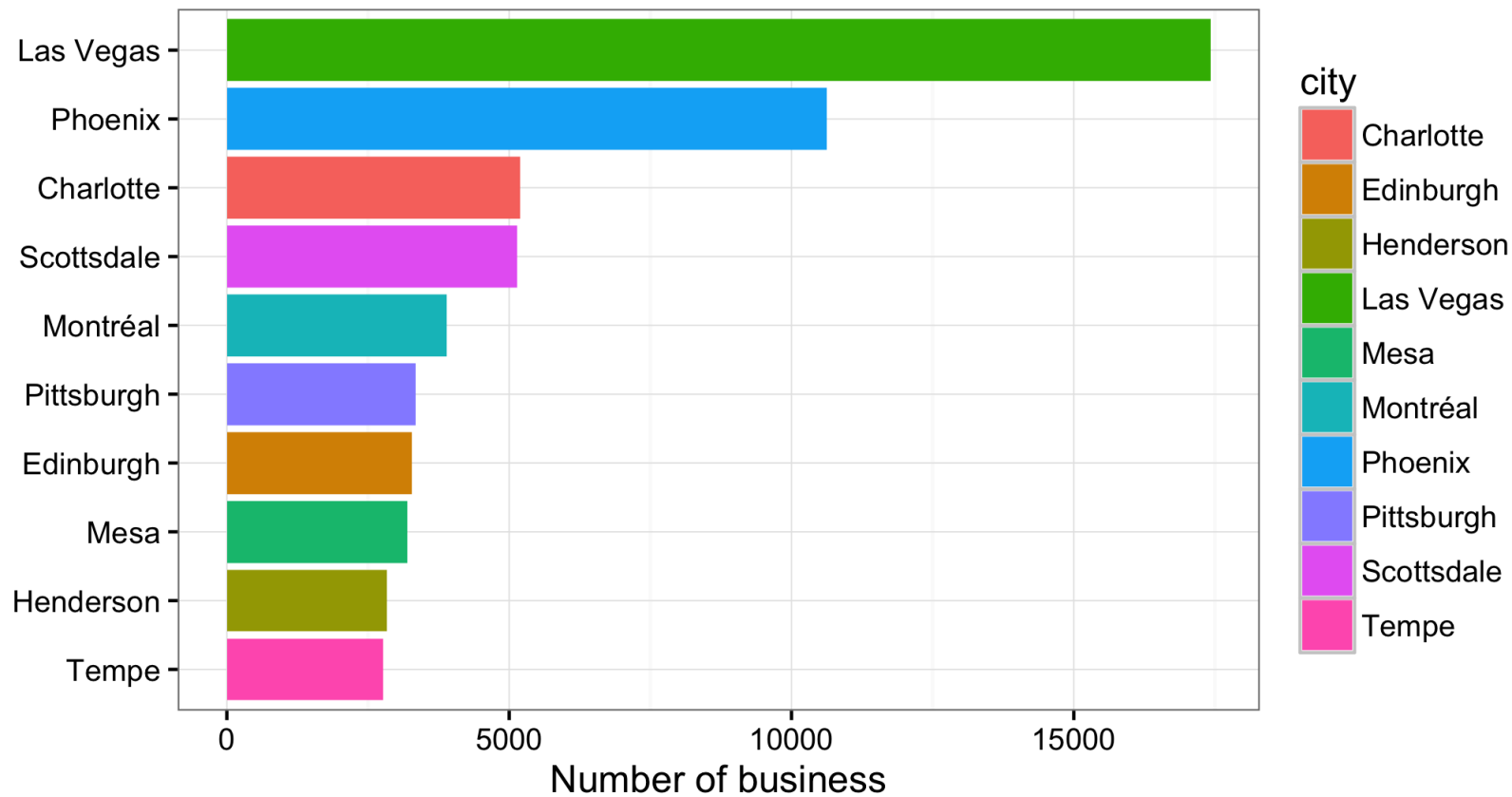
Business Statistics

Average number of reviews vs ratings



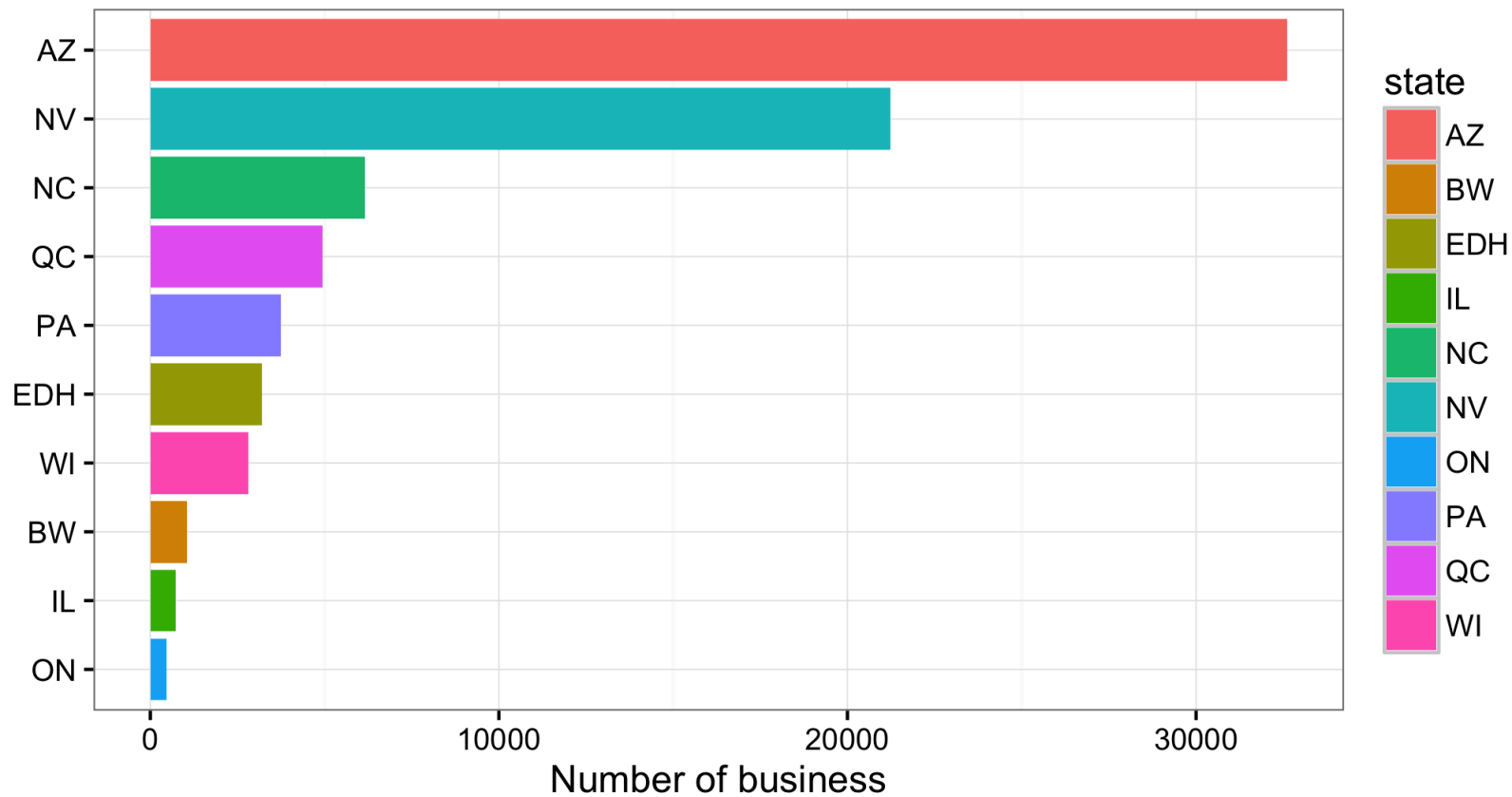
Business Statistics

Top 10 Cities in terms of number of business



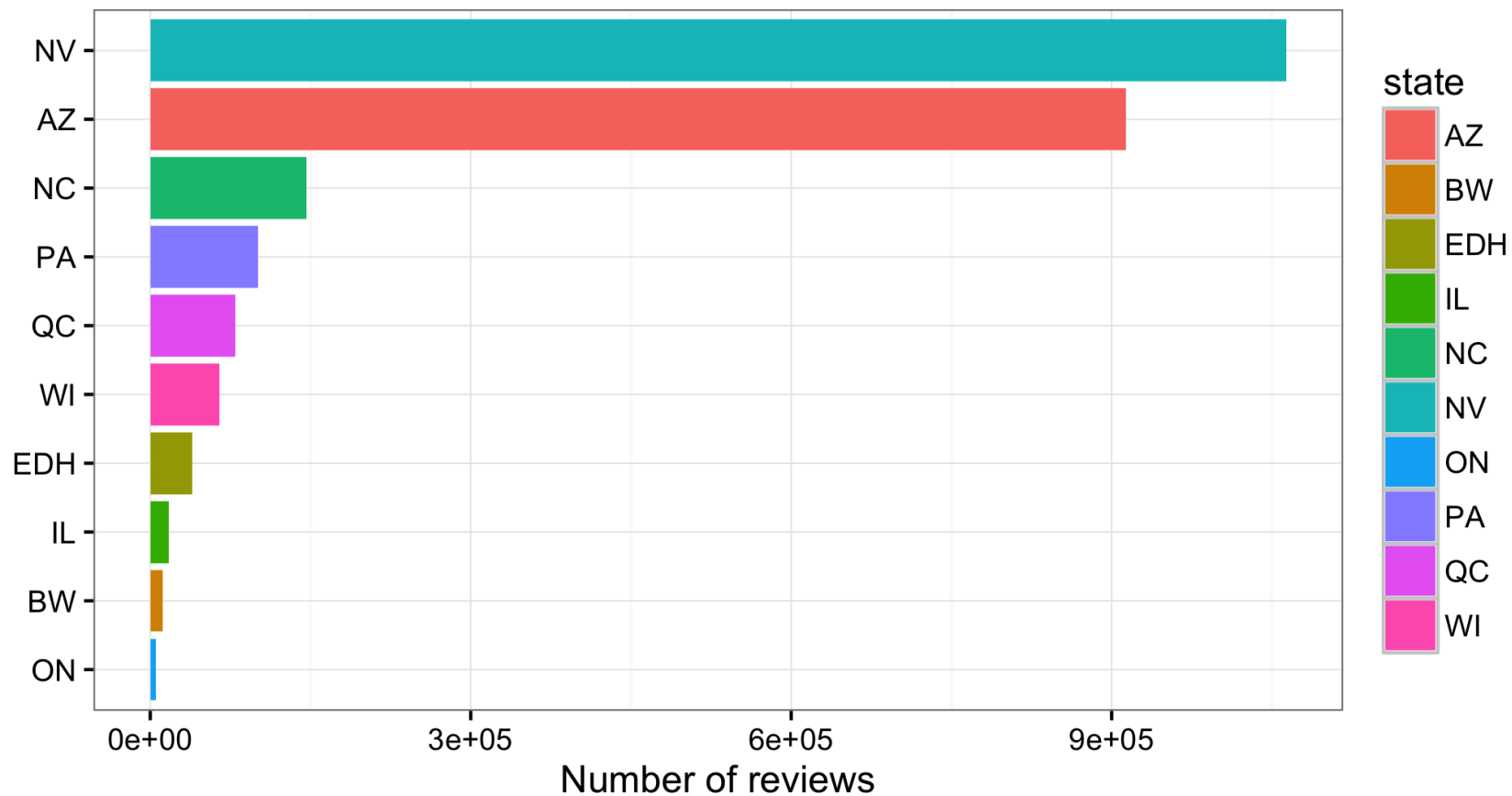
Business Statistics

Top 10 States in terms of number of business



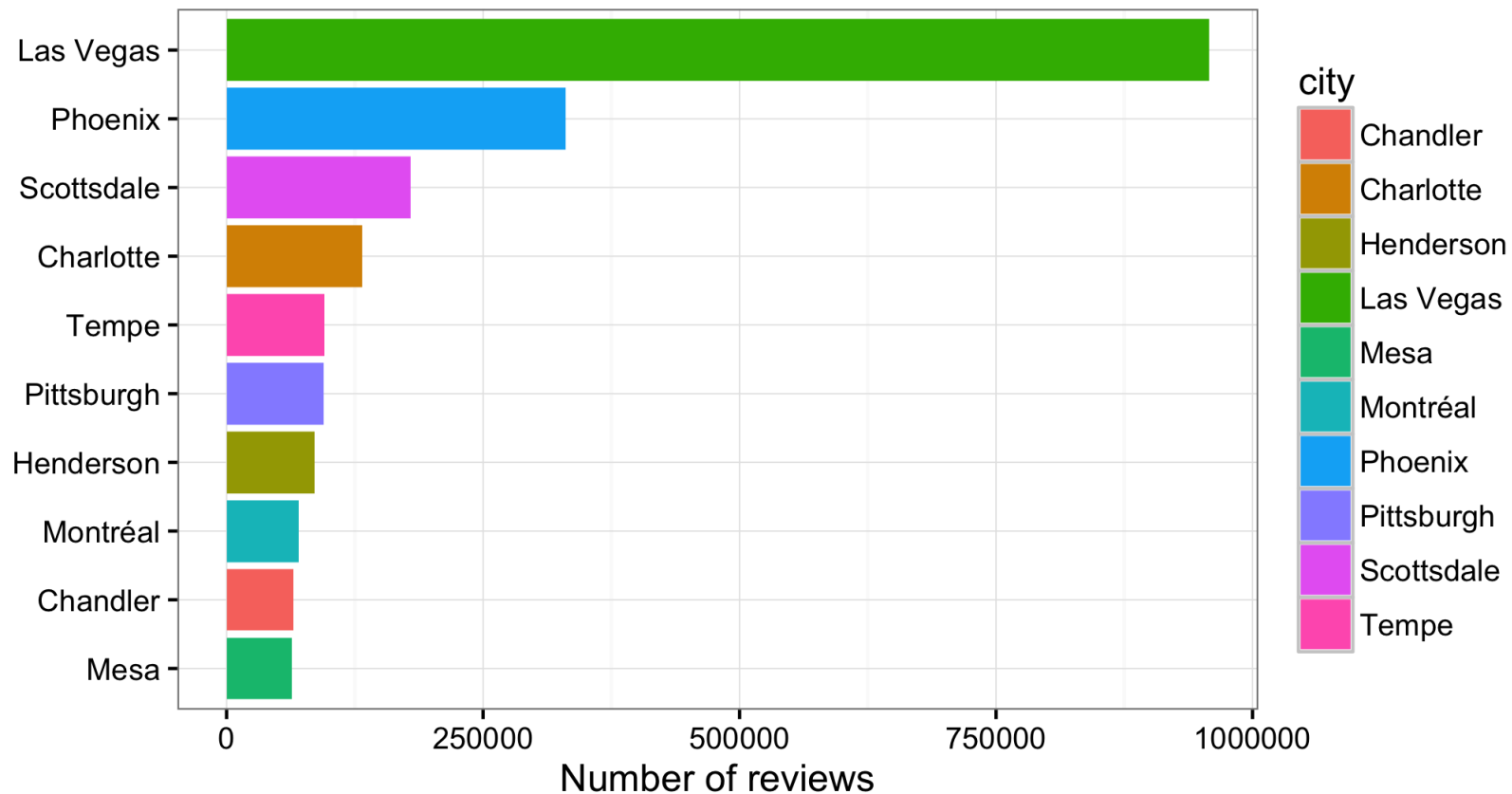
Business Statistics

Top 10 States in terms of number of reviews



Business Statistics

Top 10 cities in terms of number of reviews



Thanks !