Yelp Reviews: Las Vegas

Jeremy Mednik, Stephen Owen, Karim Naous

Objectives

What we wanted to learn…

- ► Unique Lemmas Between Restaurant Types

- ► Unique lemmas between review ratings

- ► General Information about the companies
  - ► Geographical relationships
  - ► Average review ratings

- ► Grouping comparisons together because of their similar features

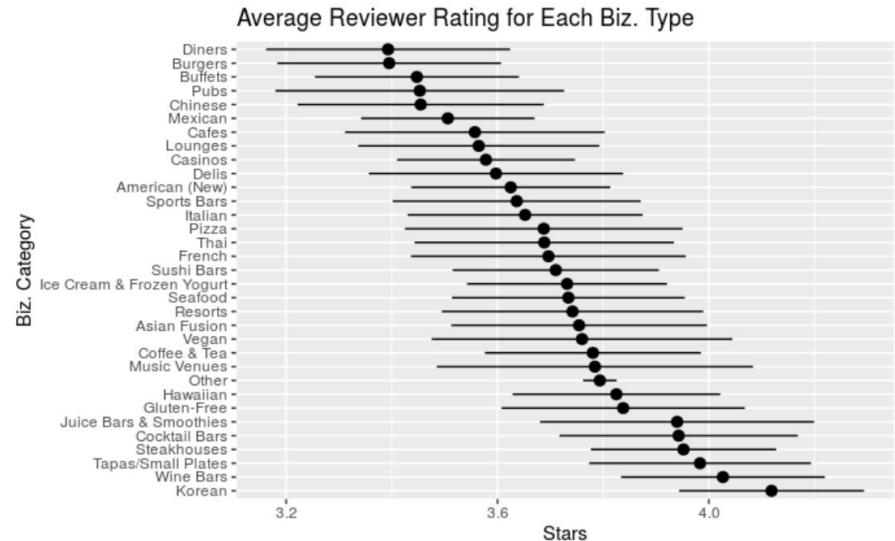# Business Type General Information

Highest Average Rating:
► Korean Restaurants

Lowest Average Rating
► Diners

Casinos are among the lower tier of average ratings for restaurants.

Music Venues is among the higher tier of average ratings for restaurants and has a high standard error spread.

Average Reviewer Rating for Each Biz. Type

Clusters

```
[1]  "Ice Cream & Frozen Yogurt; Juice Bars & Smoothies"
[2]  "Coffee & Tea; Cafes"
[3]  "Thai; Chinese; Vegan"
[4]  "Music Venues; Lounges"
[5]  "Italian; Delis; Gluten-Free; Pizza"
[6]  "Other; Mexican; Hawaiian"
[7]  "Seafood; Steakhouses; Tapas/Small Plates; Cocktail Bars; Wine Bars; French"
[8]  "Sports Bars; Pubs"
[9]  "Resorts; Casinos"
[10] "Diners; Burgers; American (New)"
[11] "Buffets; Korean"
[12] "Sushi Bars; Asian Fusion"
```

12 Clusters

Distinct comparisons between companies.

Interesting observations:
- ► Buffets and Korean are placed in the same cluster, suggesting that maybe Korean restaurants in Las Vegas have a buffet option.
- ► Casinos and Resorts are placed in the same cluster, which makes sense since they are not primarily restaurants.

Lemmas

Found 5 lemmas for each business type

Most lemmas make sense:
► "Mein" for Chinese, "Caffeine" for Cafes, "Cheeseburger" for Burgers

Unusual lemmas were present in our findings:
► "Rabbit" for Seafood, "Naked" for Sports Bars, "808" for Asian Fusion, "Daddy" for Pubs

Casino lemmas:
► "Tower", "Slot", "Alley", "Bowling", "Orleans"

	doc_id	tokens
1	American (New)	ihop; Shack; Benedict; Go; bull
2	Asian Fusion	808; eel; marrow; Kalbi; Hamachi
3	Buffets	Buffet; carving; ayce; mongolian; bulgogi
4	Burgers	Cheeseburger; tat; Smash; Wendy; McDonalds
5	Cafes	latte; caffeine; Croissant; Mocha; toffee
6	Casinos	Tower; slot; alley; bowling; Orleans
7	Chinese	mein; China; Won; Go; wonton
8	Cocktail Bars	marrow; Crème; libation; Marrow; Bang
9	Coffee & Tea	Starbuck; caffeine; Mocha; Starbucks; starbuck
10	Delis	sicilian; deli; Port; Cannoli; Deli
11	Diners	Diner; Metro; Hangover; grit; Denny
12	French	macaron; escargot; Croissant; ex; Mon
13	Gluten-Free	Blaze; Greens; custom; Kale; Bison
14	Hawaiian	katsu; tonkatsu; Hawaiian; Moco; Plate
15	Ice Cream & Frozen Yogurt	Shack; yogurt; Queen; Custard; Yogurtland
16	Italian	Cheeseburger; Mile; Miracle; Puck; fennel
17	Juice Bars & Smoothies	smoothie; bulgogi; Rachel; yogurt; boba
18	Korean	Kalbi; bulgogi; banchan; kimchi; Korea
19	Lounges	Lounge; duel; fruity; piano; smokey
20	Mexican	Chipotle; carnita; Carne; verde; carne
21	Music Venues	guitar; Hard; stage; dancer; Elvis
22	Other	shoe; appointment; tire; cupcake; golf
23	Pizza	Knights; Arena; National; Seattle; practice
24	Pubs	PT; Daddy; IPA; nacho; pint
25	Resorts	Carlo; suite; Monte; Encore; Mirage
26	Seafood	-The; chimichanga; Oyster; rabbit; Tamales
27	Sports Bars	Tavern; Naked; tavern; Mexico; sand
28	Steakhouses	hibachi; Georges; Steakhouse; Jean; steakhouse
29	Sushi Bars	yellowtail; Nigiri; handroll; toro; ayce
30	Tapas/Small Plates	Carson; puree; Tivoli; falafel; fennel
31	Thai	Pad; Panang; satay; pad; thai
32	Vegan	Vegan; Pad; Trattoria; Uni; Kurobuta
33	Wine Bars	Parma; Panang; Raku; Carnevino; collection

Which lemmas are tied to positive reviews vs. negative reviews for casinos?

Let us look at lemmas most predictive for **Casinos** and number of **Stars** together:

```
      doc_id                                                              tokens
1 Casinos                shit; holiday; lasagna; palm; 100x; 50x; 9; absolute; albertson's; ashamed
2 Casinos GCB; miso; oyster; unpredictable; checkbook; Fitz; relationship; volatile; Best; bingo
3 Casinos    Java; grand; Sunset; comp; Starbuck; Orlean; box; neighborhood; traditional; willing
4 Casinos                pie; pork; Barley; band; Eiffel; serious; Paris; *; admit; affordable
5 Casinos        Sauce; \\uuu/; Portofino; request; duck; fight; roasted; treat; artist; barkeep
```

A negative Casino review may mention: ***Lasagna, miso, oyster***.

A neutral Casino review may mention: ***Starbucks, java***.

A positive Casino review may mention: ***duck, roasted, treat, barkeep, pie, pork, barley***.

Let us compare casinos to their most similar business category.

The closest neighbor to **Casinos**, using angle distances between documents as the measure, is: **Resorts.**

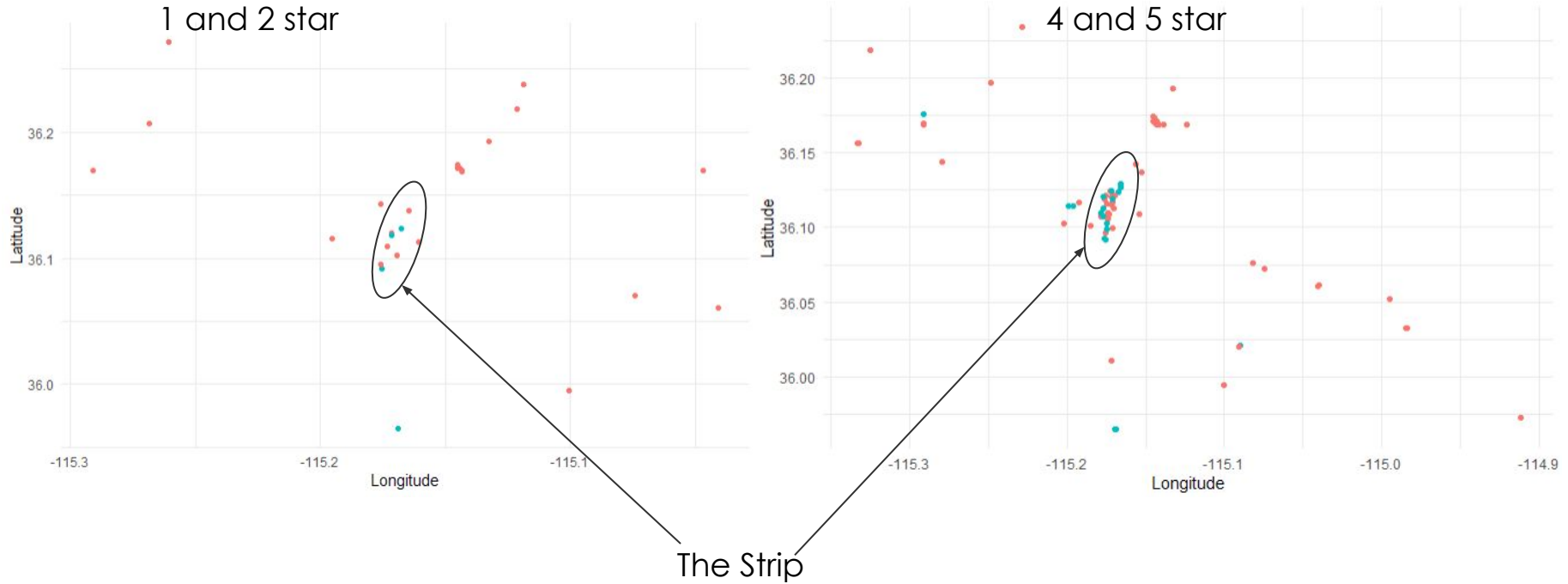Let us look at lemmas most predictive for **Resorts** and number of **Stars** together:

```
      doc_id                                                              tokens
1 Resorts Blv; hate; nightmare; park; S; staircase; building; direction; finally; lock
2 Resorts          meh; rather; sushi; 100; anything; confusing; shop; 17; 6th; accident
3 Resorts    ..; Butter; bathroom; hope; desk; bartender; big; Chocolate; couple; front
4 Resorts                        she; dish; spa; doughnut; head; luck; spicy; *; !; day
5 Resorts        Vdara; ;; amazing; art; bike; noise; oh; service; absolute; cappuccino
```

A negative review for a resort may mention: *sushi.*

A neutral review for a resort may mention: *butter, bartender, chocolate*.
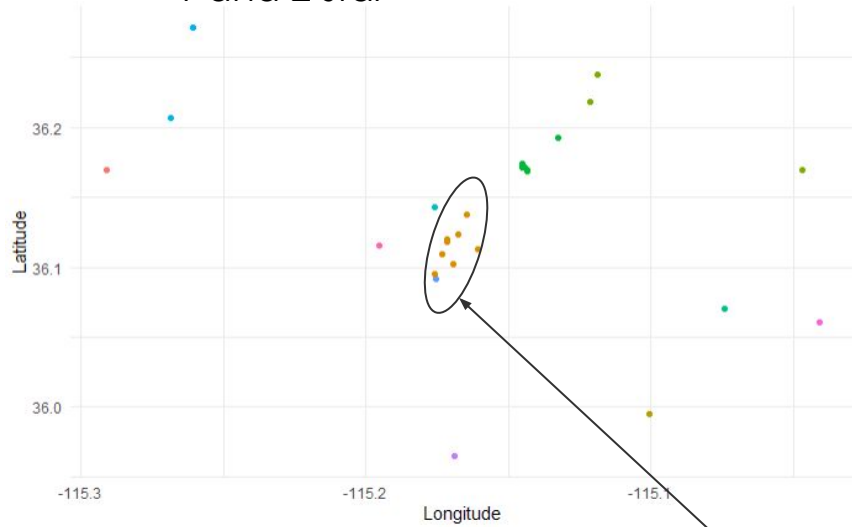
A positive review for a resort may mention: *dish, doughnut, spicy, cappuccino*.
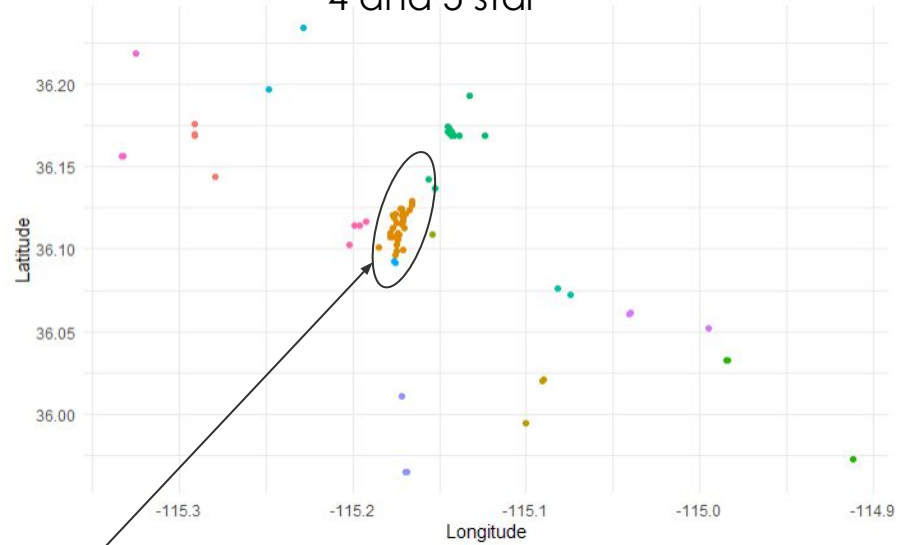
Casinos and Resorts: Locations

1 and 2 star

4 and 5 star

The Strip

Casinos and Resorts: Clustering

1 and 2 star

4 and 5 star

The Strip

Conclusion

- Clustering can provide us extremely interesting insight into people's thoughts regarding various locations.
- It can also help us make informed decisions without a large amount of prior knowledge

But most of all...

- If you're going to go to Las Vegas, stay at a resort on the strip, order duck for dinner, and have coffee & donuts for dessert.