**Problem Statement by Goldman Sachs:**

**Team name: PMSR**

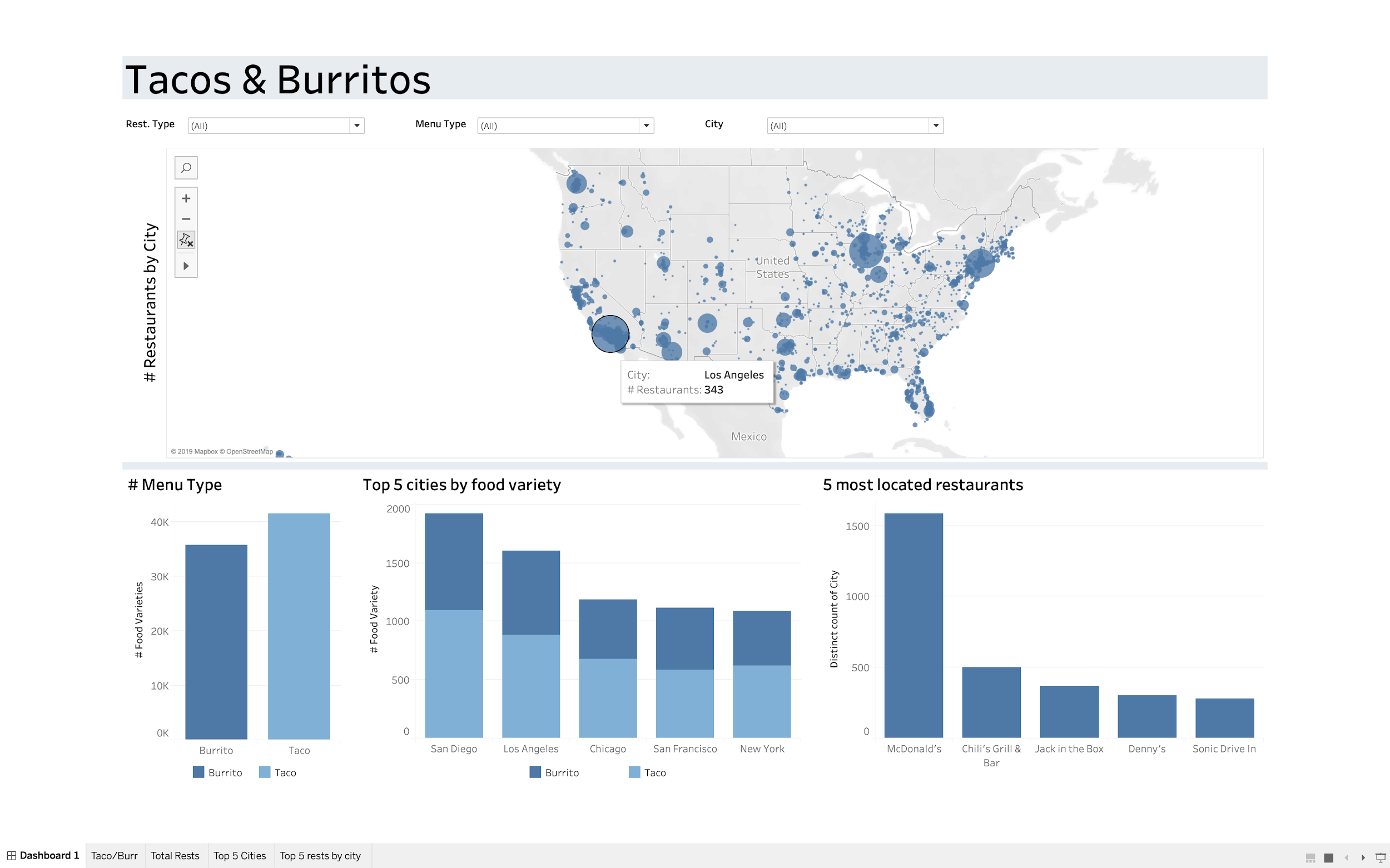
**Team Members: Prathiksha, Monika, Sandeep, Roline**

Data processing:

1. The dish variety is determined as Taco or Burrito from the dataset column ‘menus.name’. This information is further used for Taco vs Burrito comparisons. Case-insensitivity is considered while deriving the above information.
2. Mexican vs. Non-Mexican restaurant distinction is done based on the ‘categories’ column of the given dataset. The presence of the word “Mexican Restaurant” and “Tex-Mex Restaurant” in the “categories” field is used to distinguish between Mexican and non-Mexican restaurants.
3. Contents of the dataset column ‘menus.description’ is parsed and analyzed to get ingredient information.
4. Data is cleaned w.r.t case sensitive words and apostrophe. For example, McDonald’s, Mcdonalds and McDonalds are all changed to McDonald’s.
5. The ‘NaN’ values present in 'menus.description' field of the dataset are eliminated to draw a better word cloud.

Tools/libraries used:

1. Wordcloud
2. Pandas
3. Matplotlib
4. Tableau (for data visualization)
5. Regex
6. Numpy
7. **Which city in US has the highest number of restaurants(selling Tacos and Burritos)?**
8. **Which US city sells the highest variety of Tacos and Burritos?**
9. **Which are the most common restaurants selling Tacos and Burritos in US?**
10. **Which item has more variety? Tacos or Burritos?**
11. **Which regions in the US have more restaurants selling Tacos and Burritos?**



1. Restaurant count is visualized in the above map, with the size of the circular marker. We see major cities namely Los Angeles, Chicago have more such restaurants
2. San Diego, Los Angeles, Chicago, San Francisco and New York are top 5 cities with the maximum variety of Tacos and Burritos.
3. We see that Mc Donalds has highest branches in the US, with more than 1500 restaurants spread across different cities followed by ‘Chilli’s grill & bar’, ‘Jack in the box’, ‘Denny’s and ‘Sonic Drive in’.
4. As we see from the leftmost bar chart, Taco’s have more varieties than Burritos.
5. We see that restaurants are generally concentrated in North-eastern, Southern border and west coast. Tacos and Burritos do not seem very popular in the Northern and Central regions and US.
6. **For Mexican restaurants, are there more varieties of Tacos or Burritos available?**
7. **Which Mexican restaurants are located in maximum number of cities?**
8. **Cities with highest number of Mexican restaurants?**

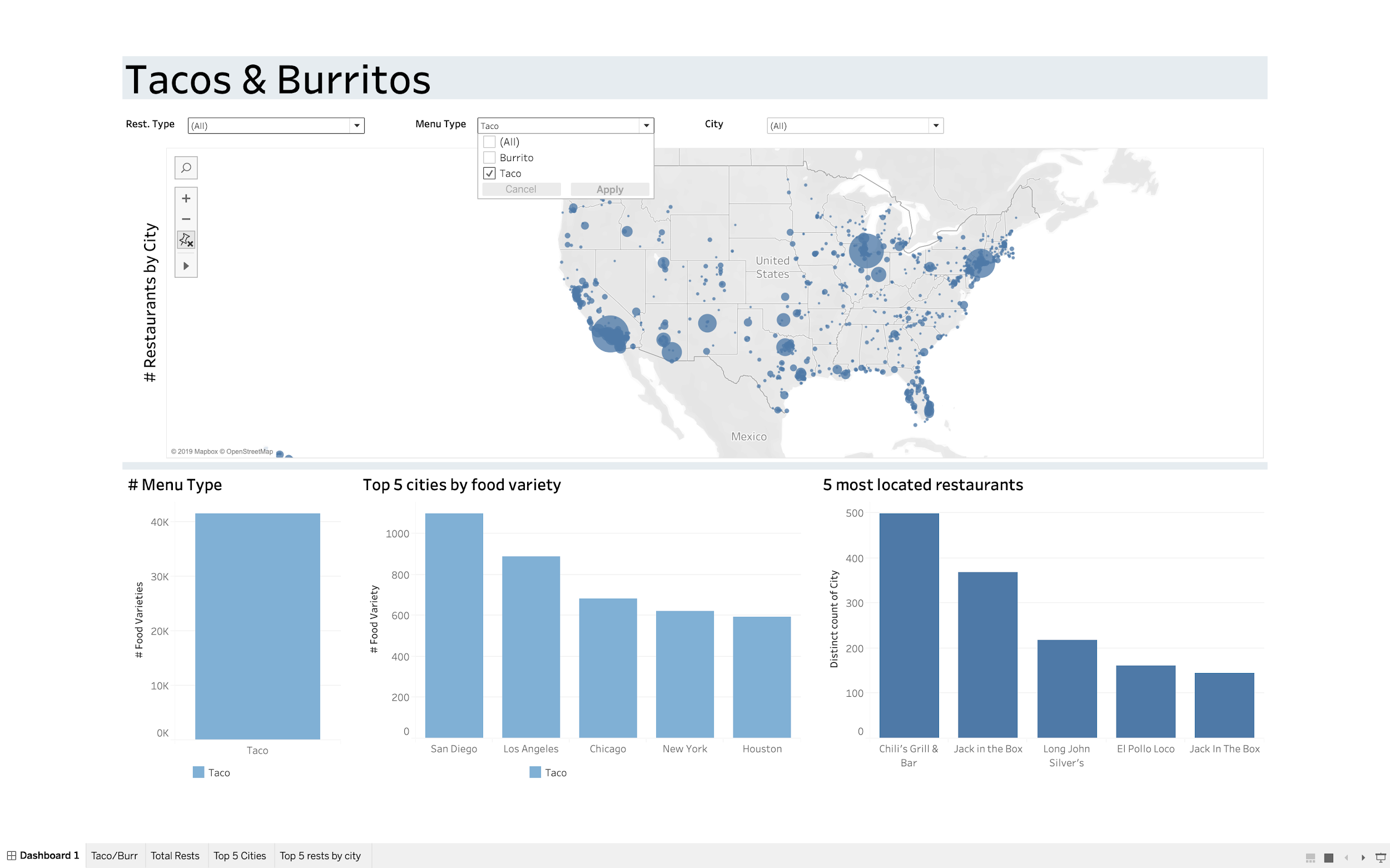


5. From the above graphs, we can see that there are more varieties of Taco dishes in **Mexican restaurants.**

6. Mexican restaurants that have outlets in maximum number of cities/regions are Chili’s Grill and Bar, El Pollo Loco and Taco Bell.

7. Mexican restaurants are more in number in San Diego followed by Los Angeles.

1. **Which city is best city for Taco lovers and has maximum varieties for Tacos available?**
2. **Which city is the best city for Burrito lovers and has maximum varieties for Burritos available?**



9. Maximum varieties for Tacos are found in San Diego followed by Los Angeles.

10. Similar results are seen for Burritos. And hence San Diego sells the highest variety of Tacos and Burritos.

**Word clouds for analyzing the ingredients in tacos and burritos:**

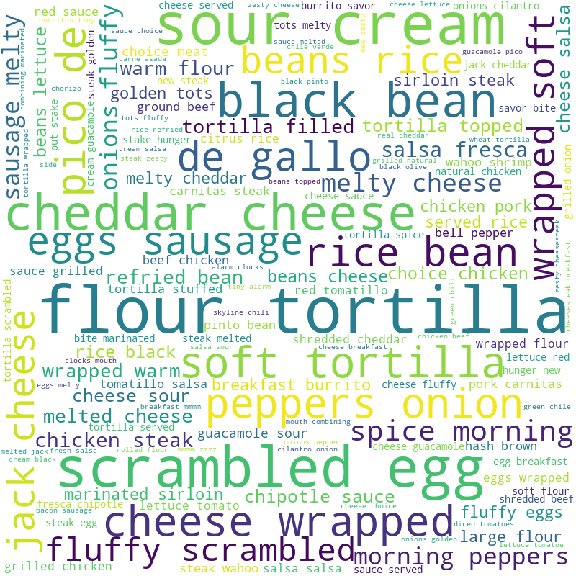
We created a word cloud in python for ingredients in Tacos and Burritos. We also created separate word clouds for visualizing ingredients used in Mexican and non-Mexican restaurants.

1. **Which are the common ingredients used in Tacos in US ? What insights we can derive from that?**



Here, we can see that most of the tacos are made with corn tortilla followed by flour tortilla. Pico de gallo, lettuce, cheese (cheddar, American,etc.), beef, sour cream, taco sauce, crunchy corn are also common ingredients of taco. A lot of words like spicy, seasoned, lime, etc. show that restaurants prefer adding spices and flavors in the tacos and they might be preferred by customers. We see words like crisp, crunchy, etc indicating that tacos texture.

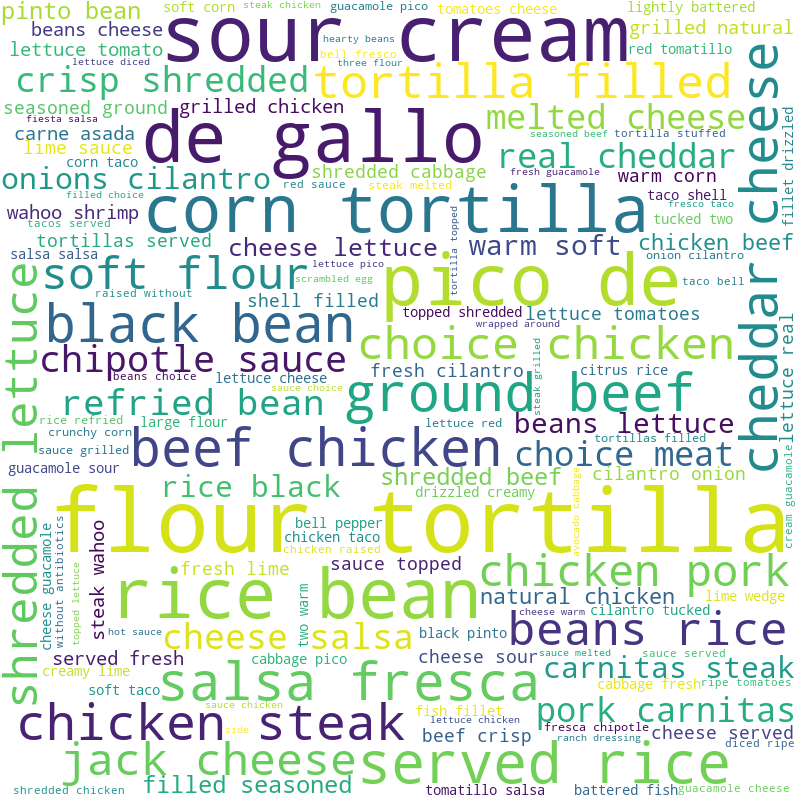
1. **Which are the common ingredients used in Burritos in US ? What insights we can derive from that?**



From the above word cloud, we can conclude that most burritos are made up of flour tortillas.

Also, we see main ingredients like sour cream, black bean, pico de gallo, cheese (cheddar, jack,etc.), eggs, sausage, rice, peppers, onions, etc. A lot of words like melted, fluffy, soft, etc. indicate that burritos are generally preferred in soft texture.

1. **What ingredients are generally used in Mexican restaurants in US ? What insights can we derive from them ?**



Here, we see the common ingredients of tacos and burritos made in Mexican restaurants are sour cream, pico de gallo, tortilla (corn and flour), black beans, beef, chicken, pork, rice and beans, salsa, Fresca, chipotle etc. These look like authentic Mexican items because they are not that commonly used in non-Mexican restaurants (seen from the next word cloud). There is also a good balance of carbohydrates and protein ingredients.

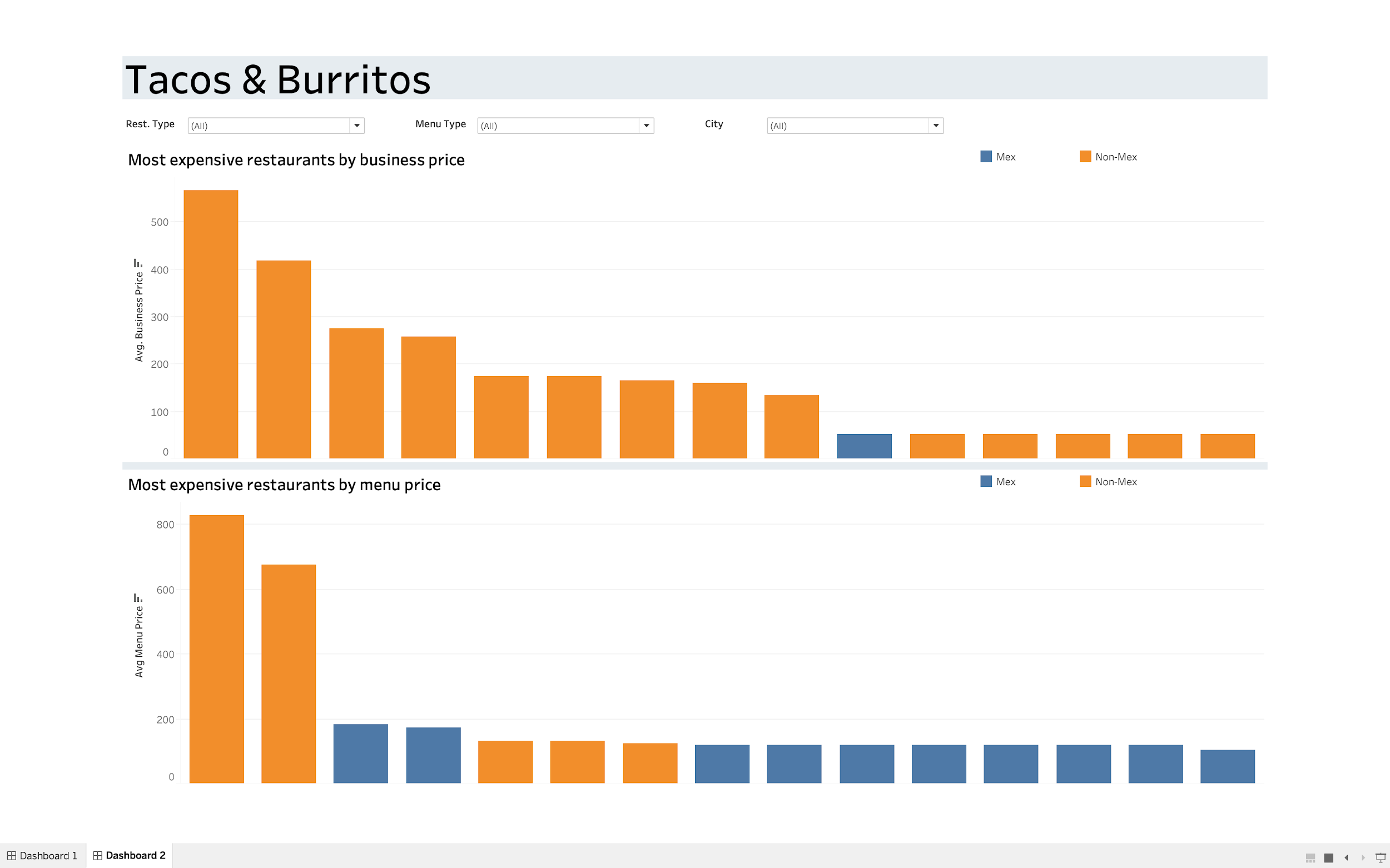
1. **What ingredients are generally used in non-Mexican restaurants in Tacos and Burritos in US ? What insights can we derive from them?**



Here, we see Non-Mexican restaurants commonly use ingredients such as cheddar cheese, scrambled eggs, lettuce etc in preparing Tacos and Burrito. They do use basic Mexican items like pico de gallo, sour cream, etc. We see that Non-Mexican restaurants like using cheddar, cheese, cream flavored tacos and Burittoes whereas in Mexican restaurants meat(Non-veg) taste is preferred.

**14. What type of restaurant is more expensive with respect to business price?**

**15. What type of restaurant is more expensive with respect to menu item price?**

****

1. **Non-Mexican restaurants are expensive when the business price is considered.**
2. **The menu item price for few Non-mexican restaurants is quite high, but we see an even distribution of mexican and non-mexican restaurants thereafter unlike the business price, where non-mexican restaurants are predominant.**