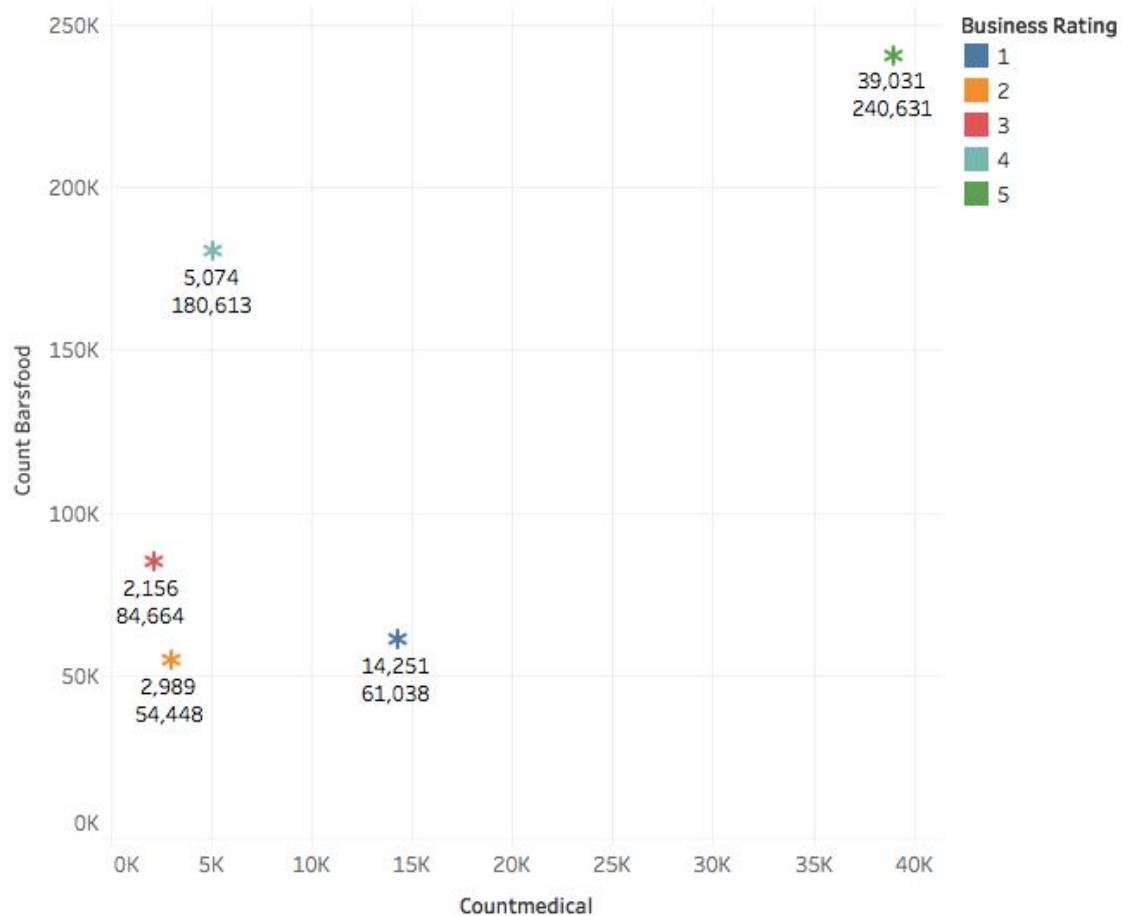


Shivangi Singh
Final Visualization Project

Visualization 1

Sheet 1



Countmedical vs. Count Barsfood. Color shows details about Business Rating. The marks are labeled by Countmedical and Count Barsfood.

TREND SHOWN:

The visualization represents the total number of ratings for businesses for a each given rating (1,2,3,4,5) for the category labelled food, medical

We can see that people review hospitals less than they review food places, this might be due to two reasons.

- 1) Less people visit hospital
- 2) There are less hospitals than food joints.

The general trend is that people are most likely to give a rating of 5 in both the medical and food units.

Whereas the least likely given rating is different for food and medical units being 2 in food and 3 in the medical unit.

SQL Query Executed:

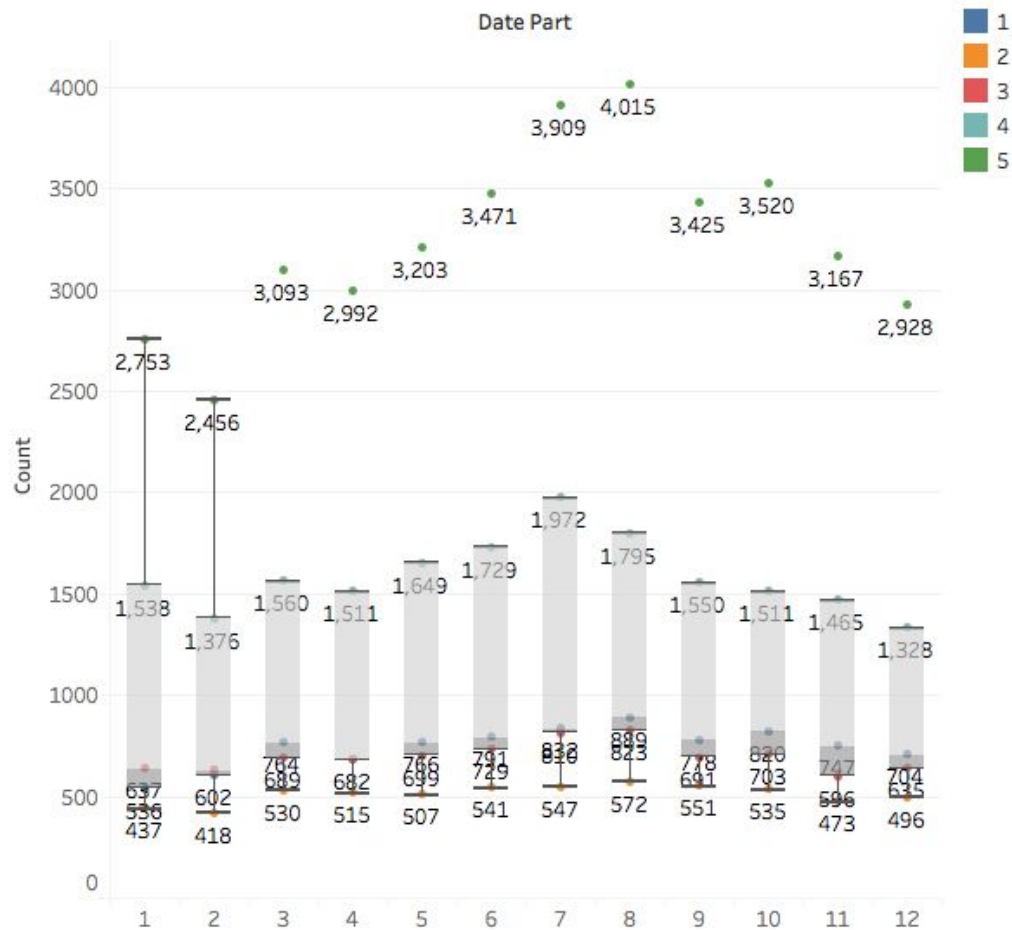
```
CREATE VIEW bars_food AS
SELECT count(*) AS count_barsfood,
       r.business_rating
FROM business b,
     review r
WHERE (((b.business_categories)::jsonb ? 'Bars'::text) OR ((b.business_categories)::jsonb ?
'Food'::text)) AND (r.business_id = b.business_id))
GROUP BY r.business_rating
ORDER BY r.business_rating;
```

```
CREATE VIEW mediingeneral AS
SELECT count(*) AS countmedical,
       r.business_rating
FROM business b,
     review r
WHERE (((b.business_categories)::jsonb ? 'Health & Medical'::text) AND (r.business_id = b.business_id))
GROUP BY r.business_rating
ORDER BY r.business_rating;
```

```
SELECT b.count_barsfood,b.business_rating,m.countmedical
FROM bars_food b
INNER JOIN mediingeneral m
ON m.business_rating=b.business_rating;
```

Visualization 2

Sheet 1



Count for each Date Part. Color shows details about Business Rating.

TREND SHOWN:

The visualization represents the total number of ratings for businesses for a each given rating (1,2,3,4,5) for the category labelled food.

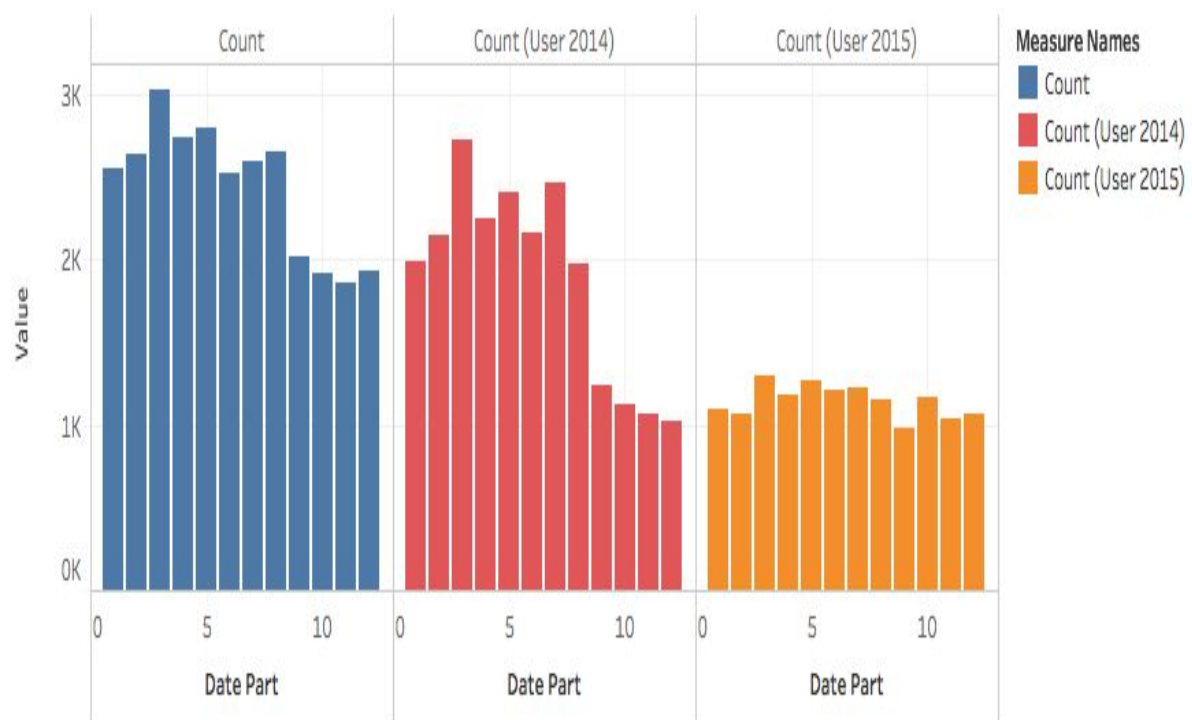
The data for the year 2015 is extracted- represented by the month of the year, the mean, media labelled in the gray scale. The general trend is that people are most likely to give a rating of 5, whereas the least likely given rating is 2. People are most active in July and August (which are the summer months in the US), and least active in December-January-February, although we can see that Jan is comparatively high this trend might be due to new year's celebrations.

SQL Query Executed:

```
SELECT count(*),r.business_rating,extract(month from r.review_date)
FROM business b,
review r
WHERE ((b.business_categories)::jsonb ? 'Food'::text) AND (r.business_id = b.business_id and
(extract(year from r.review_date):: INT ='2015'))
GROUP BY extract(month from r.review_date),r.business_rating
ORDER BY r.business_rating, extract(month from r.review_date)
```

Visualization 3

Sheet 1



Date Part vs. Count, Count (User 2014) and Count (User 2015). Color shows details about Count, Count (User 2014) and Count (User 2015). The data is filtered on Count, Count (User 2014) and Date Part. The Count filter ranges from 1853 to 3033. The Count (User 2014) filter ranges from 1030 to 2729. The Date Part filter ranges from 1 to 12 and keeps Null values.

TREND SHOWN:

In this visualization I have tried to compare the tip activity over 2013, 2014, 2015 for users who joined yelp in 2012. I have also divided the year by the month so that we can also see how the tip activity has changed not only over the years but also the months.

We can see that the users were most actively giving tips during the first year of them joining yelp. The activity has declined over the years. When the users were pretty active in tipping most tips were given during March, and comparatively less tips were given in the last three month of the year.

SQL Query Executed:

```
CREATE VIEW users_2012 AS
SELECT *
FROM "user" u
WHERE (date_part('year'::text, u.yelping_since) = '2012'::double precision)

CREATE VIEW user_2013 AS
SELECT count(*) AS count,
       date_part('month'::text, t.tip_date) AS date_part
FROM users_2012 u,
     tip t
WHERE ((date_part('year'::text, t.tip_date) = '2013'::double precision) AND (u.user_id = t.user_id))
GROUP BY (date_part('month'::text, t.tip_date))
```

```
ORDER BY (date_part('month'::text, t.tip_date))
```

```
CREATE VIEW user_2014 AS
```

```
SELECT count(*) AS count,
```

```
date_part('month'::text, t.tip_date) AS date_part
```

```
FROM users_2012 u,
```

```
tip t
```

```
WHERE ((date_part('year'::text, t.tip_date) = '2014'::double precision) AND (u.user_id = t.user_id))
```

```
GROUP BY (date_part('month'::text, t.tip_date))
```

```
ORDER BY (date_part('month'::text, t.tip_date))
```

```
CREATE VIEW user_2015 AS
```

```
SELECT count(*),extract(month from t.tip_date)
```

```
from users_2012 u, tip t
```

```
WHERE extract(YEAR from t.tip_date) = '2015' AND u.user_id=t.user_id
```

```
GROUP BY extract(month from t.tip_date)
```

```
ORDER BY extract(month from t.tip_date)
```

Then I used Tableau to perform the union of the above tables, and plot the data.