

Project Project

Jennifer Cho, Hazel Kim, Binny Lee

Below is the data set in which we will explore, tidy-up, and discover insight from. This data was found on Kaggle and we all agreed it would be interesting to explore employee opinions of some of the biggest technology companies in the industry.

```
## Parsed with column specification:
## cols(
##   X1 = col_double(),
##   company = col_character(),
##   location = col_character(),
##   dates = col_character(),
##   `job-title` = col_character(),
##   summary = col_character(),
##   pros = col_character(),
##   cons = col_character(),
##   `advice-to-mgmt` = col_character(),
##   `overall-ratings` = col_double(),
##   `work-balance-stars` = col_character(),
##   `culture-values-stars` = col_character(),
##   `carrer-opportunities-stars` = col_character(),
##   `comp-benefit-stars` = col_character(),
##   `senior-mangemnet-stars` = col_character(),
##   `helpful-count` = col_double(),
##   link = col_character()
## )

## [1] "x1" "company"
## [3] "location" "dates"
## [5] "job_title" "summary"
## [7] "pros" "cons"
## [9] "advice_to_mgmt" "overall_ratings"
## [11] "work_balance_stars" "culture_values_stars"
## [13] "carrer_opportunities_stars" "comp_benefit_stars"
## [15] "senior_mangemnet_stars" "helpful_count"
## [17] "link"

## Observations: 67,529
## Variables: 17
## $ x1 <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ...
## $ company <chr> "google", "google", "google", "goog...
## $ location <chr> "none", "Mountain View, CA", "New Y...
## $ dates <chr> "Dec 11, 2018", "Jun 21, 2013", "Ma...
## $ job_title <chr> "Current Employee - Anonymous Emplo...
## $ summary <chr> "Best Company to work for", "Moving...
## $ pros <chr> "People are smart and friendly", "1...
## $ cons <chr> "Bureaucracy is slowing things down...
## $ advice_to_mgmt <chr> "none", "1) Don't dismiss emotional...
## $ overall_ratings <dbl> 5, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, ...
## $ work_balance_stars <chr> "4.0", "2.0", "5.0", "2.0", "5.0", ...
## $ culture_values_stars <chr> "5.0", "3.0", "4.0", "5.0", "5.0", ...
## $ carrer_opportunities_stars <chr> "5.0", "3.0", "5.0", "5.0", "5.0", ...
```

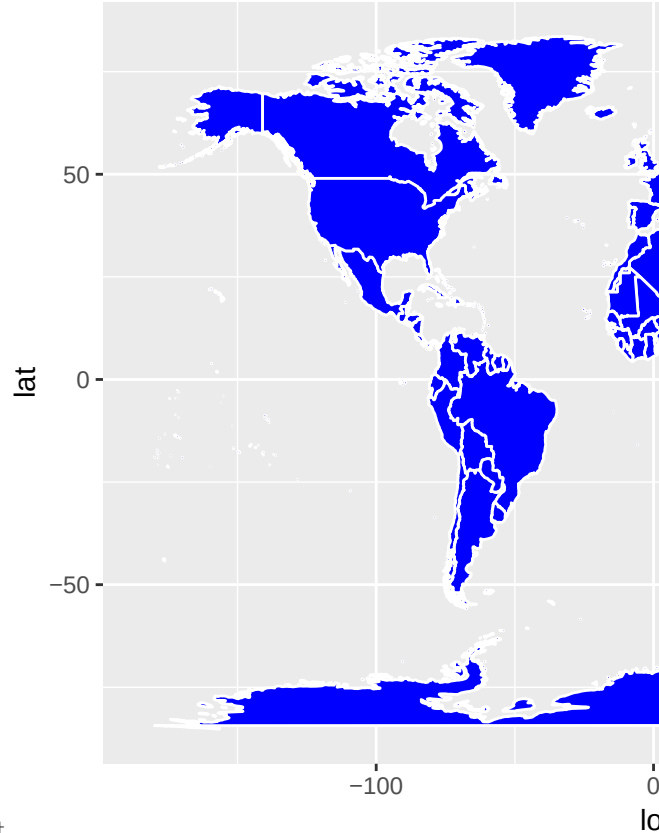
```
## $ comp_benefit_stars      <chr> "4.0", "5.0", "5.0", "4.0", "5.0", ...
## $ senior_mangemnet_stars  <chr> "5.0", "3.0", "4.0", "5.0", "5.0", ...
## $ helpful_count          <dbl> 0, 2094, 949, 498, 49, 1, 0, 0, 0, ...
## $ link                    <chr> "https://www.glassdoor.com/Reviews/...
```

The variables we are interested in using are selected below. We noticed that there are variables types that do not correlate to their actual values. For example, the columns ‘work-balance-stars’, ‘culture-values-stars’, ‘carrer-opportunities-stars’, ‘comp-benefit-stars’, and ‘senior-management-stars’ are character variables but it should be doubles instead.

Number of each company reviews and number of reviews for each city

```
## # A tibble: 6 x 2
## # Groups:   company [6]
##   company      n
##   <chr>      <int>
## 1 amazon    26430
## 2 apple     12950
## 3 facebook   1590
## 4 google     7819
## 5 microsoft 17930
## 6 netflix     810
```

Data visualization starts here

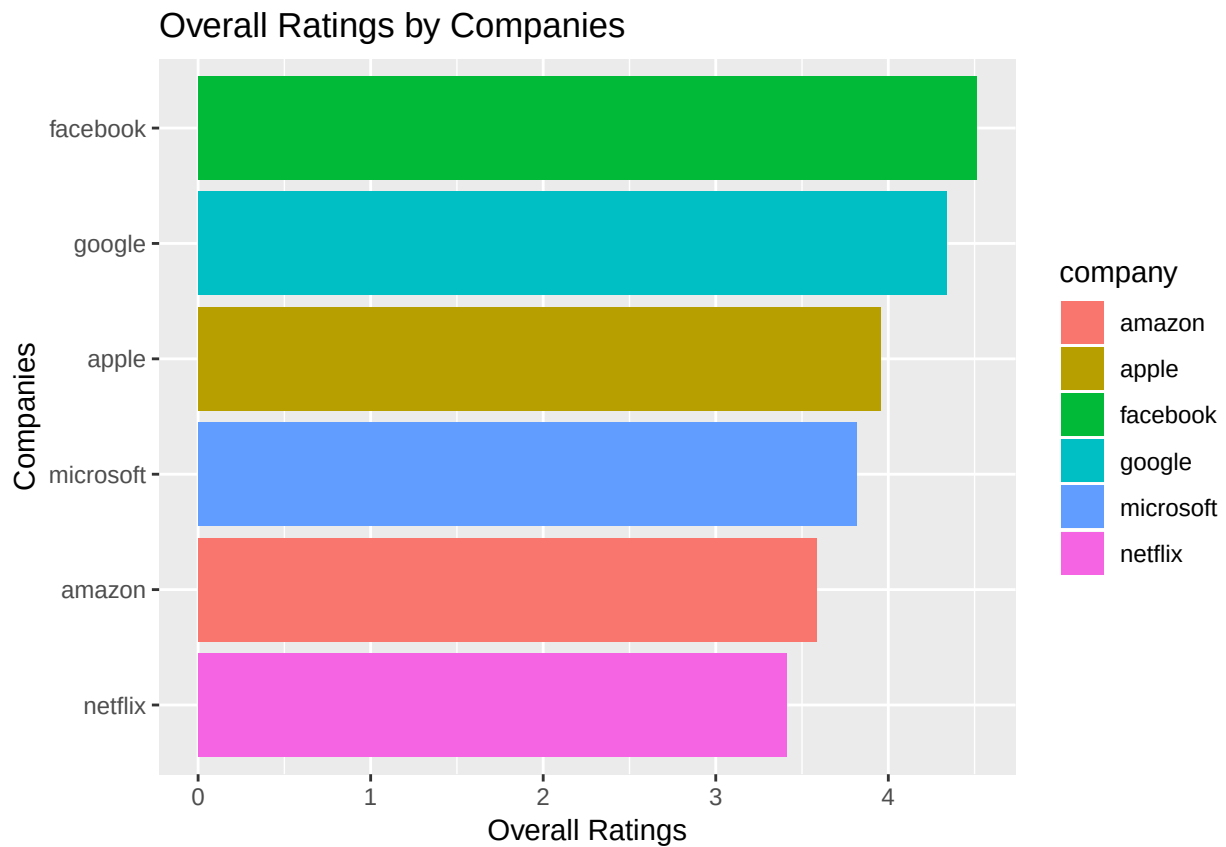


Create a map of all the locations of the companies in the review dataset

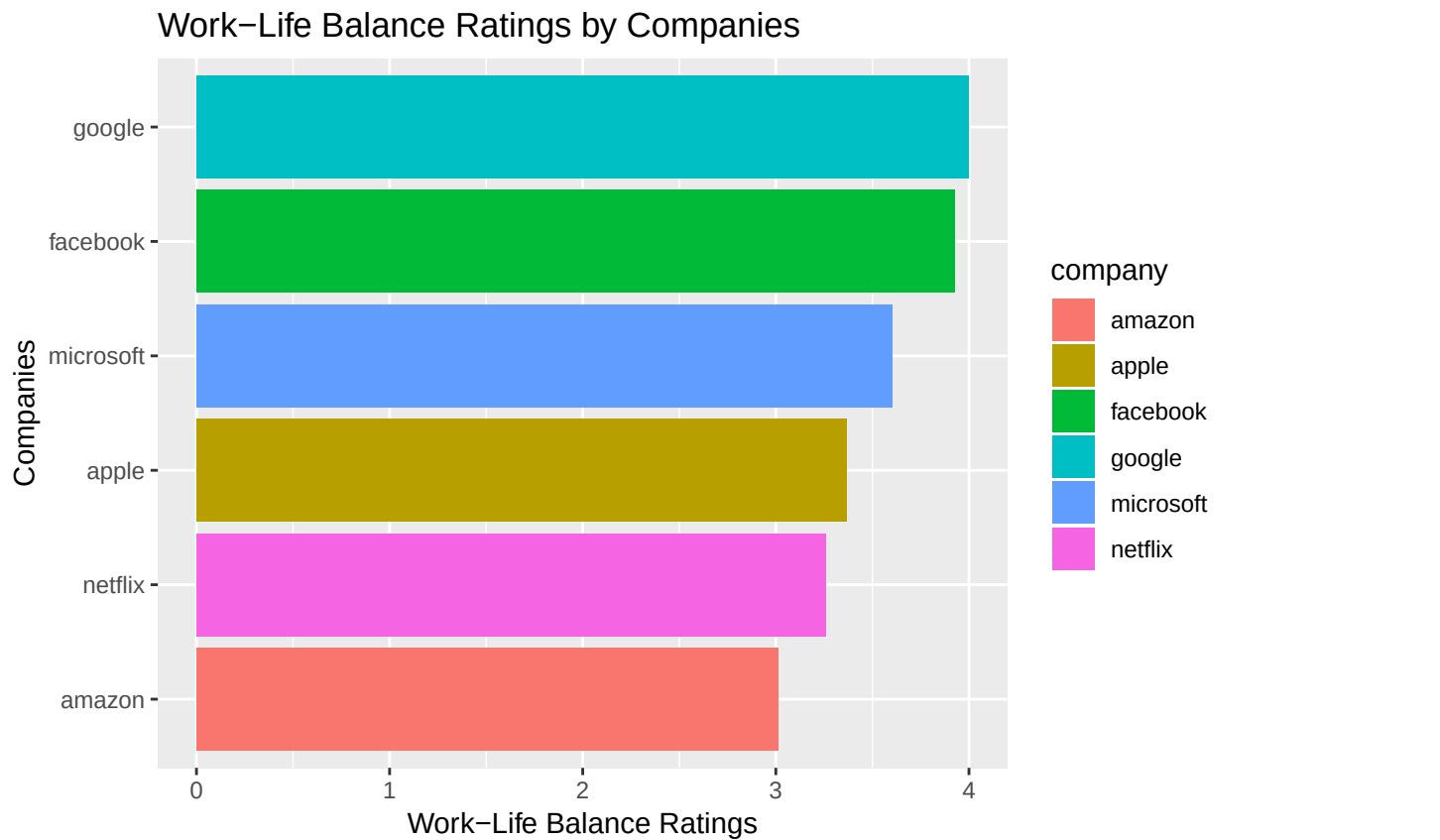
General Overview of Employee Review

Different ratings by companies

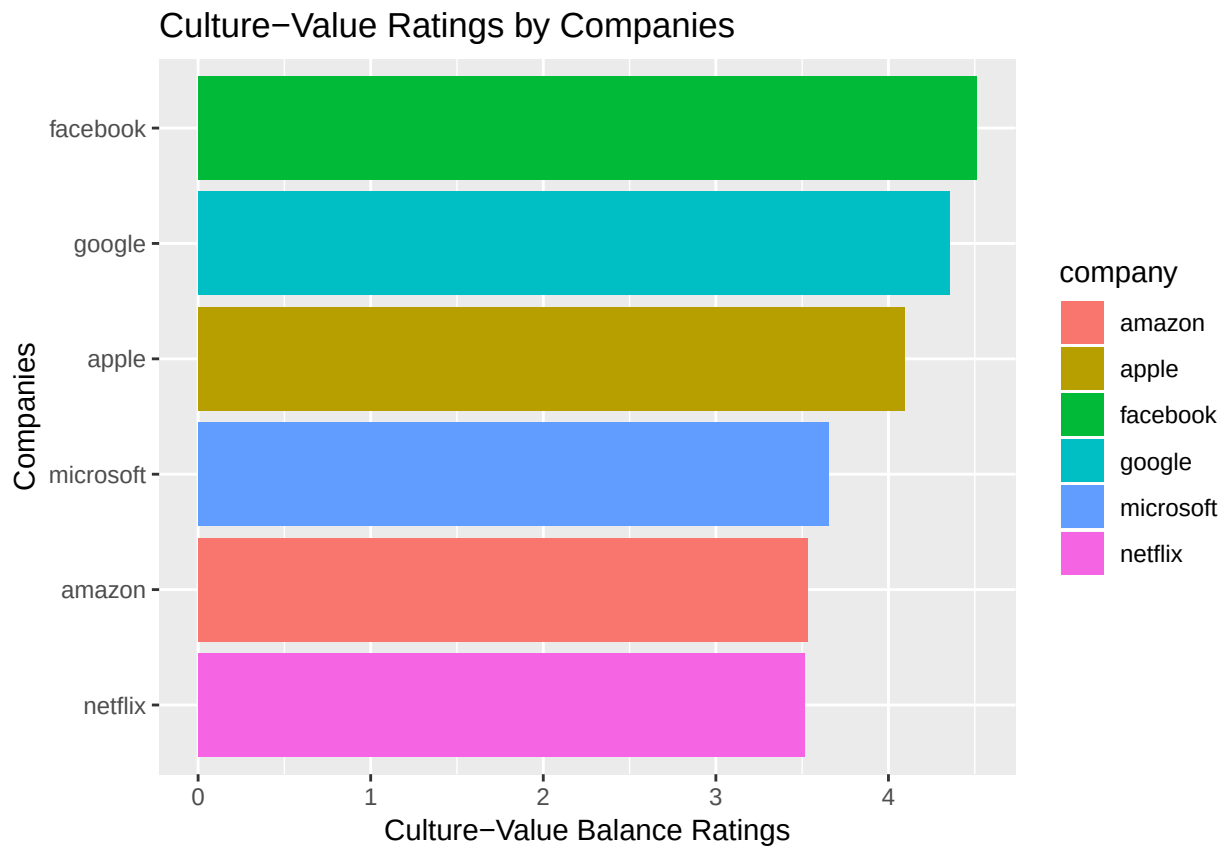
```
# overall
rev %>%
  select(company, overall_ratings) %>%
  group_by(company) %>%
  summarise(mean_overall_ratings = mean(overall_ratings)) %>%
  ggplot(aes(x = reorder(company, mean_overall_ratings), y = mean_overall_ratings, fill = company)) +
  geom_col() +
  coord_flip() +
  labs(title = "Overall Ratings by Companies", y = "Overall Ratings", x = "Companies")
```



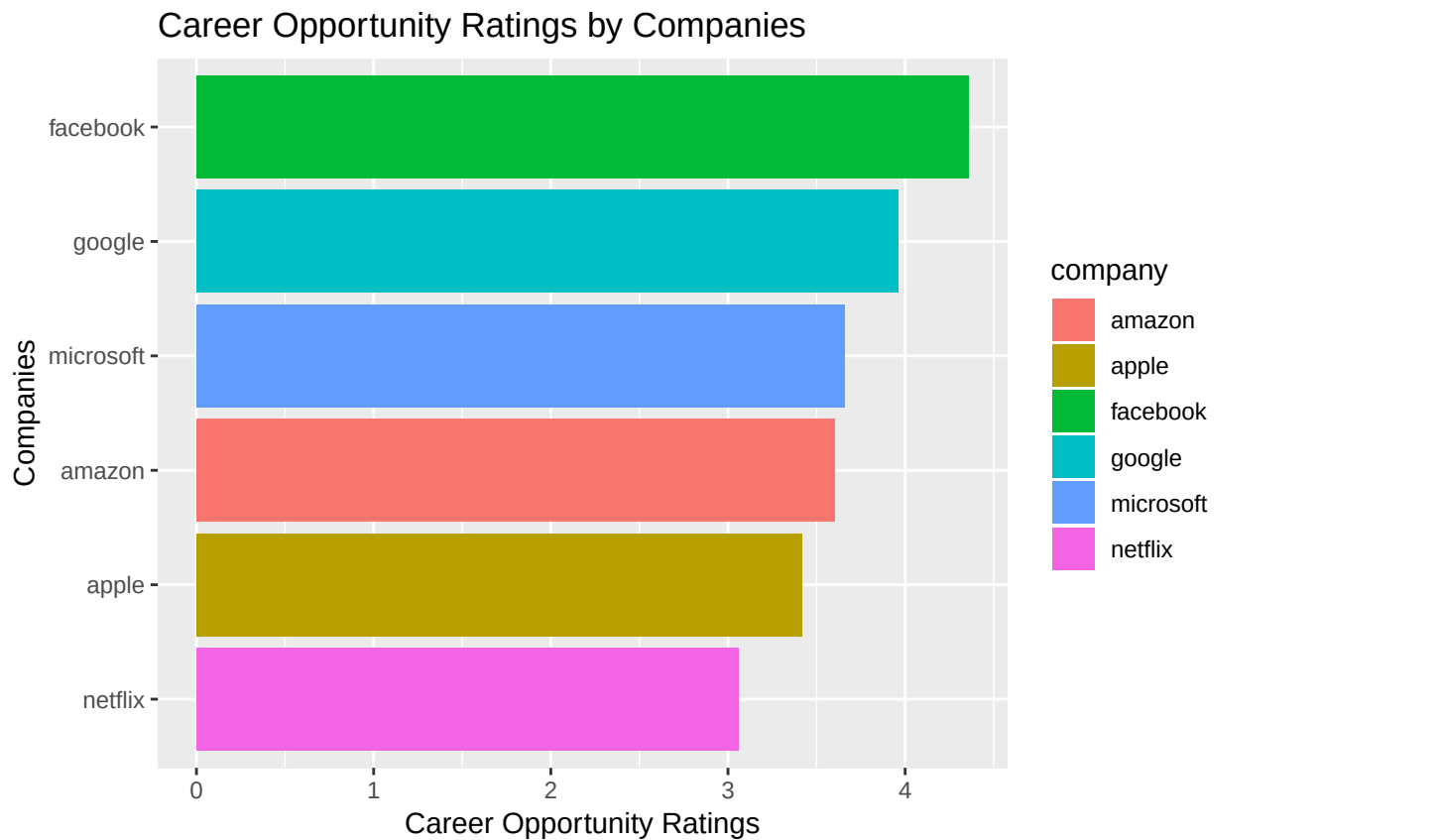
```
# work balance
rev %>%
  select(company, work_balance_stars) %>%
  filter(!is.na(work_balance_stars)) %>%
  group_by(company) %>%
  summarise(mean_work_balance_stars = mean(work_balance_stars)) %>%
  ggplot(aes(x = reorder(company, mean_work_balance_stars), y = mean_work_balance_stars, fill = company)) +
  geom_col() +
  coord_flip() +
  labs(title = "Work-Life Balance Ratings by Companies", y = "Work-Life Balance Ratings", x = "Companies")
```



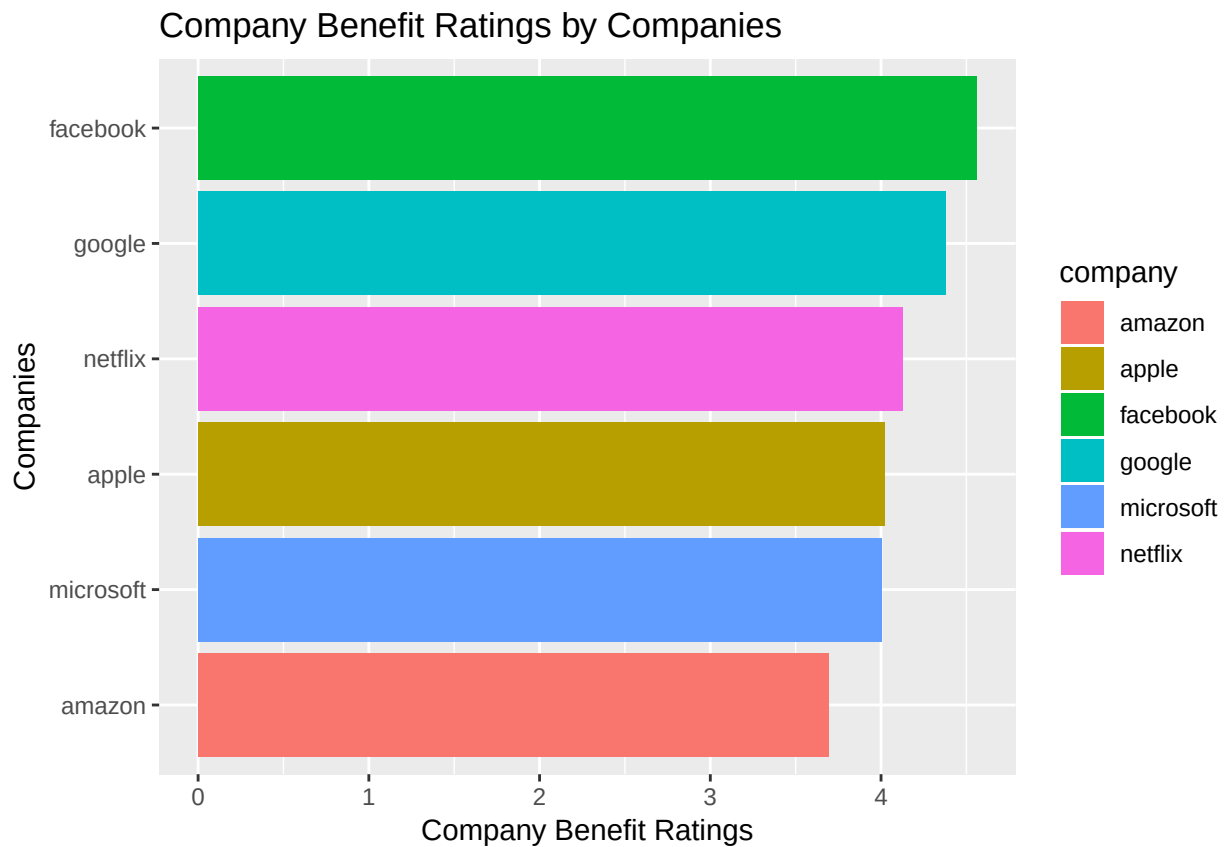
```
# culture value
rev %>%
  select(company, culture_values_stars) %>%
  filter(!is.na(culture_values_stars)) %>%
  group_by(company) %>%
  summarise(mean_culture_values_stars = mean(culture_values_stars)) %>%
  ggplot(aes(x = reorder(company, mean_culture_values_stars), y = mean_culture_values_stars, fill = company)) +
  geom_col() +
  coord_flip() +
  labs(title = "Culture-Value Ratings by Companies", y = "Culture-Value Balance Ratings", x = "Companies")
```



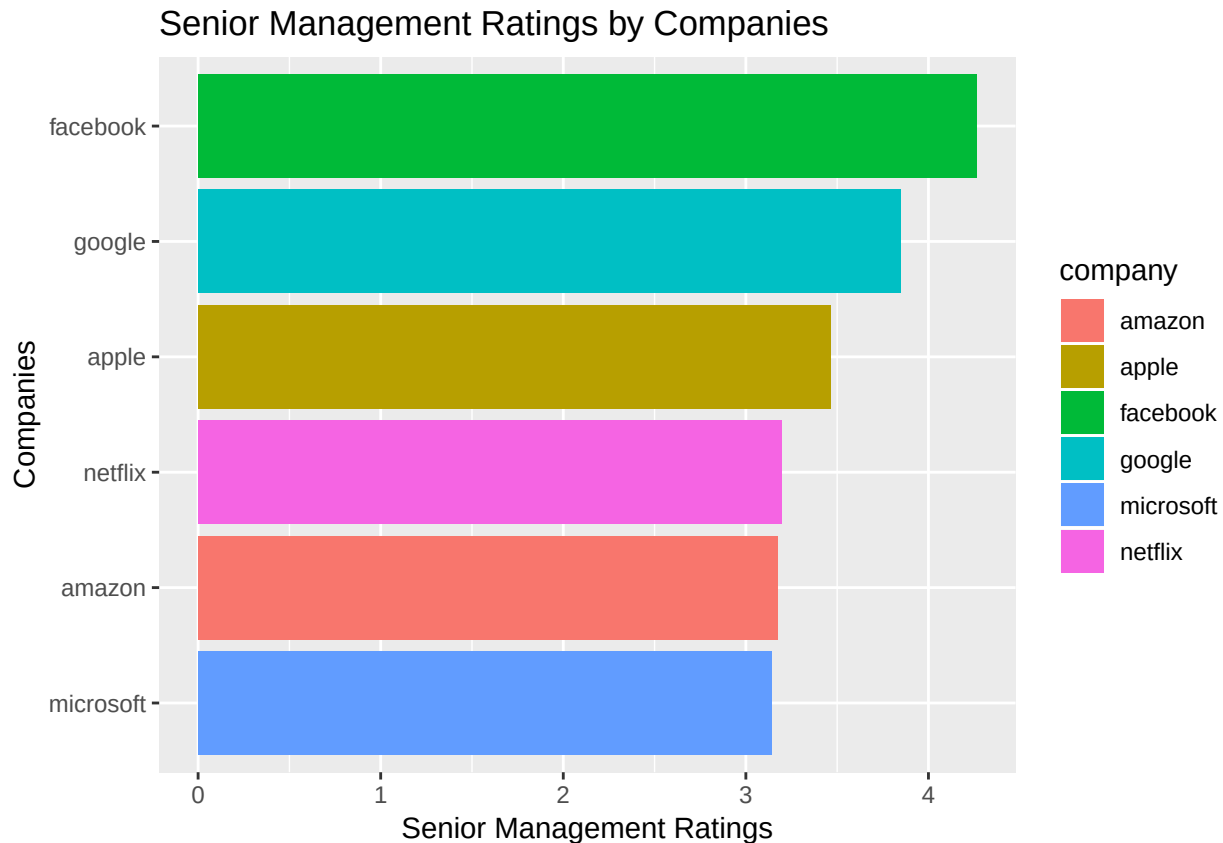
```
# career opportunity
rev %>%
  select(company, carrer_opportunities_stars) %>%
  filter(!is.na(carrer_opportunities_stars)) %>%
  group_by(company) %>%
  summarise(mean_carrer_opportunities_stars = mean(carrer_opportunities_stars)) %>%
  ggplot(aes(x = reorder(company, mean_carrer_opportunities_stars), y = mean_carrer_opportunities_stars)) +
  geom_col() +
  coord_flip() +
  labs(title = "Career Opportunity Ratings by Companies", y = "Career Opportunity Ratings", x = "Companies")
```



```
# comp benefit
rev %>%
  select(company, comp_benefit_stars) %>%
  filter(!is.na(comp_benefit_stars)) %>%
  group_by(company) %>%
  summarise(mean_comp_benefit_stars = mean(comp_benefit_stars)) %>%
  ggplot(aes(x = reorder(company, mean_comp_benefit_stars), y = mean_comp_benefit_stars, fill = company)) +
  geom_col() +
  coord_flip() +
  labs(title = "Company Benefit Ratings by Companies", y = "Company Benefit Ratings", x = "Companies")
```



```
# senior management
rev %>%
  select(company, senior_mangemnet_stars) %>%
  filter(!is.na(senior_mangemnet_stars)) %>%
  group_by(company) %>%
  summarise(mean_senior_mangemnet_stars = mean(senior_mangemnet_stars)) %>%
  ggplot(aes(x = reorder(company, mean_senior_mangemnet_stars), y = mean_senior_mangemnet_stars, fill =
  geom_col() +
  coord_flip() +
  labs(title = "Senior Management Ratings by Companies", y = "Senior Management Ratings", x = "Companies")
```

Employee Reviews by US cities

All locations of all reviews in US

```
states <- map_data("state")
#states

loc <- us.cities %>%
  separate(name, c("city", "state"), sep = " ") %>%
  filter(long >= -130)
```

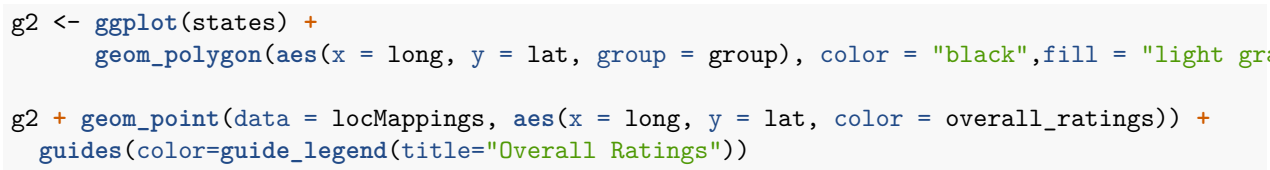
```
## Warning: Expected 2 pieces. Additional pieces discarded in 311 rows [11,
## 16, 24, 28, 29, 33, 39, 42, 54, 56, 58, 59, 66, 67, 88, 92, 93, 96, 97,
## 103, ...].
```

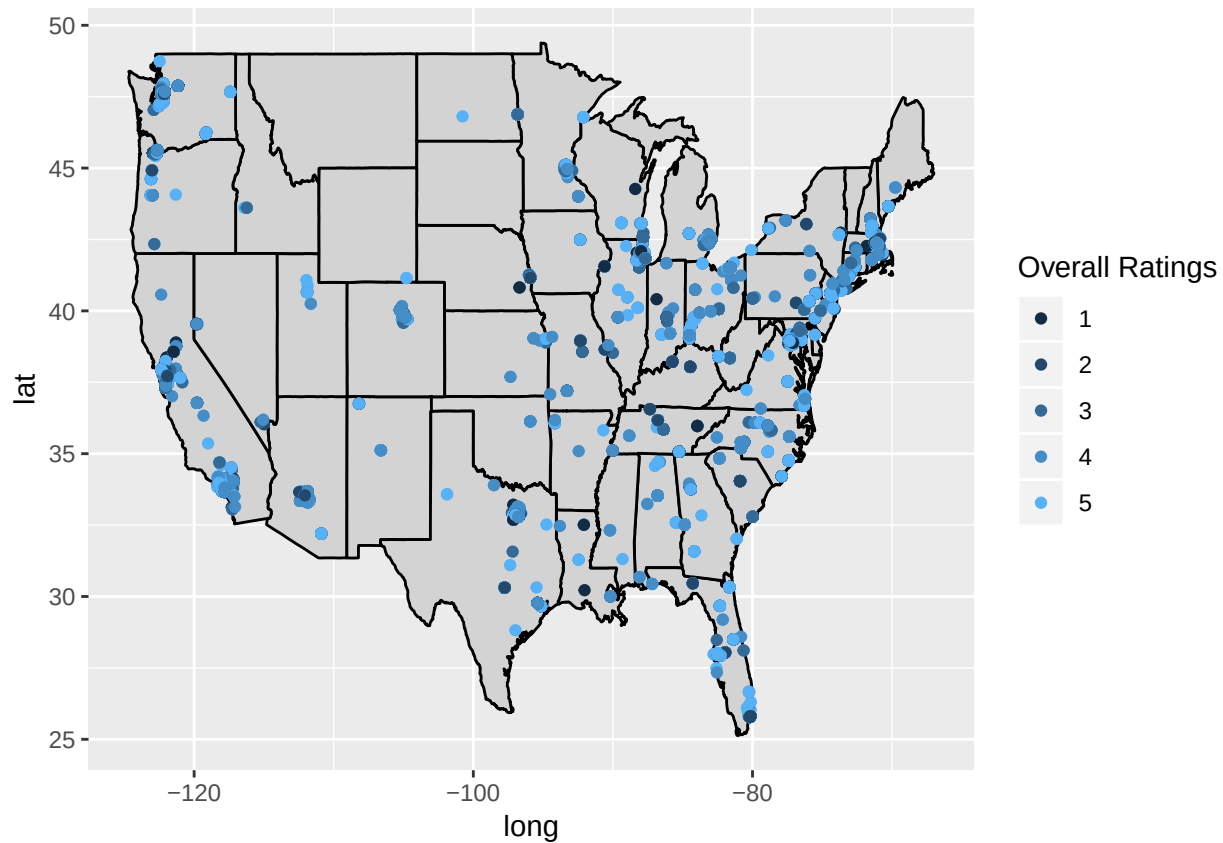
```
glimpse(loc)
```

```
## Observations: 1,001
## Variables: 7
## $ city      <chr> "Abilene", "Akron", "Alameda", "Albany", "Albany",...
## $ state     <chr> "TX", "OH", "CA", "GA", "NY", "OR", "NM", "LA", "V...
## $ country.etc <chr> "TX", "OH", "CA", "GA", "NY", "OR", "NM", "LA", "V...
## $ pop       <int> 113888, 206634, 70069, 75510, 93576, 45535, 494962...
## $ lat       <dbl> 32.45, 41.08, 37.77, 31.58, 42.67, 44.62, 35.12, 3...
## $ long      <dbl> -99.74, -81.52, -122.26, -84.18, -73.80, -123.09, ...
```

```
locMappings <- inner_join(rev, loc, by = "city")

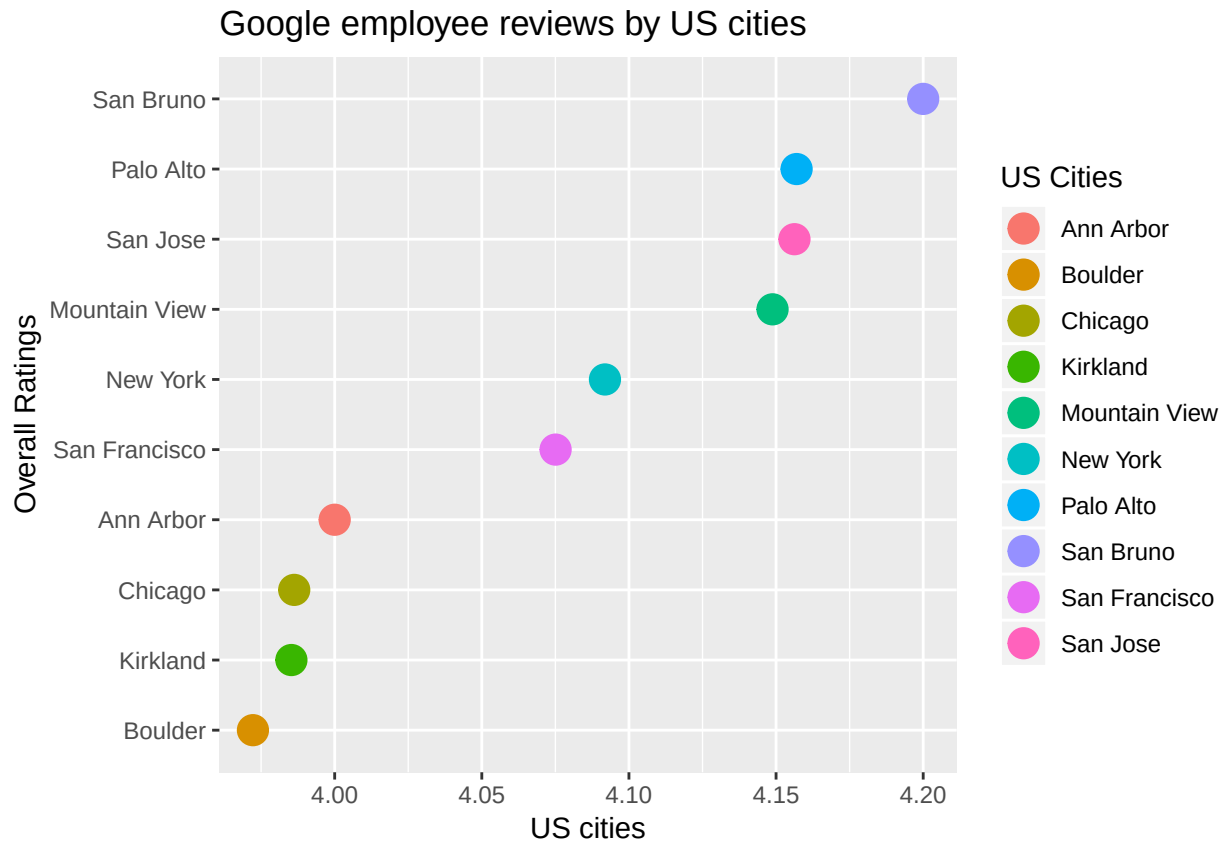
g <- ggplot(states) +
  geom_polygon(aes(x = long, y = lat, fill = region, group = group), color = "white") +
  guides(fill=FALSE)
g + geom_point(data = locMappings, aes(x = long, y = lat), color = "blue") +
  scale_size(name = "Employees")
```





Overall Rating by different cities

```
# Overall ratings in US Google
rev %>%
  filter(country == "USA") %>%
  filter(company == "google") %>%
  group_by(city) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = "city") %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  ggplot(aes(x = reorder(city, mean_overall_rating_city), y = mean_overall_rating_city, color = city)) +
  geom_point(size = 5) +
  coord_flip() +
  labs(title = "Google employee reviews by US cities", x = "Overall Ratings", y = "US cities", color = city)
```



```
# map for google
county <- map_data("county")
states <- map_data("state")

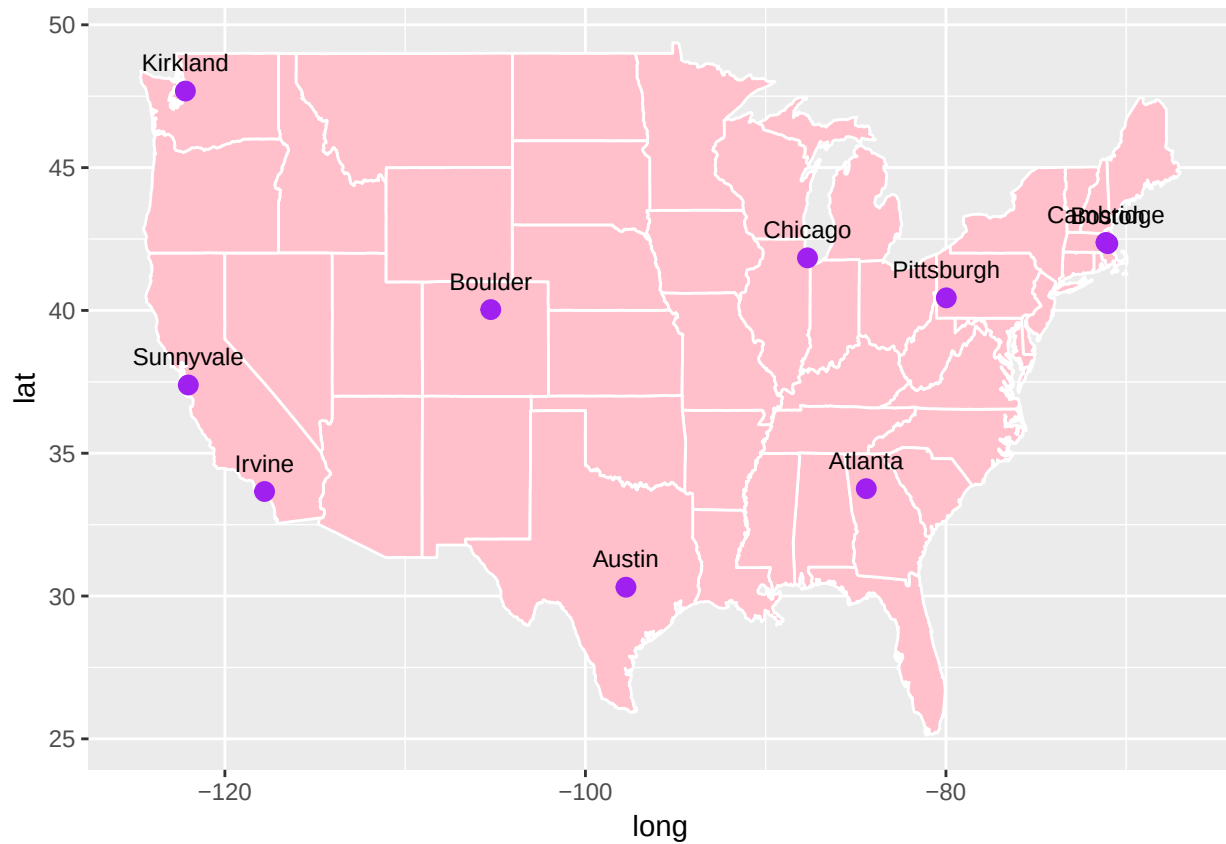
loc <- us.cities %>%
  separate(name, c("city", "state"), sep = " ") %>%
  filter(long >= -130)

## Warning: Expected 2 pieces. Additional pieces discarded in 311 rows [11,
## 16, 24, 28, 29, 33, 39, 42, 54, 56, 58, 59, 66, 67, 88, 92, 93, 96, 97,
## 103, ...].

locMappings <- inner_join(rev, loc, by = "city") %>%
  filter(country == "USA") %>%
  filter(company == "google") %>%
  group_by(city) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = "city") %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  left_join(loc, by = "city")

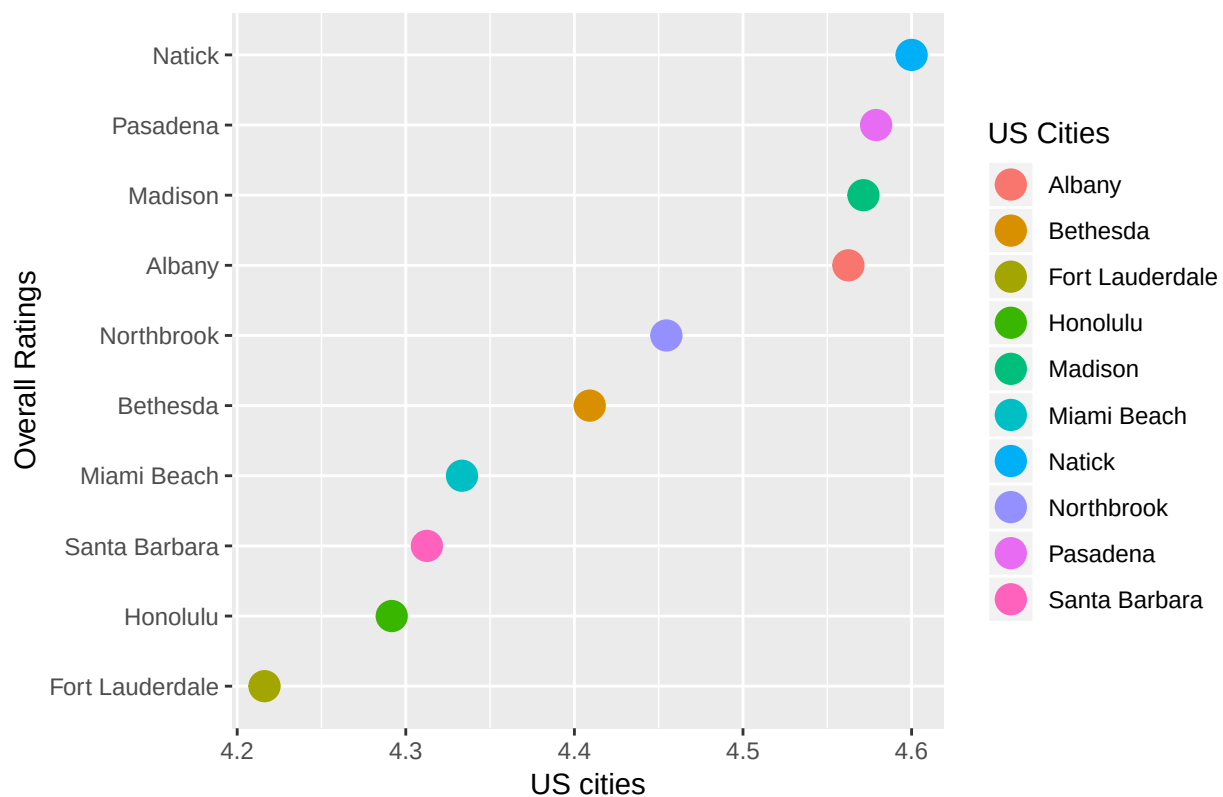
ggplot(states) +
  geom_polygon(aes(x = long, y = lat, fill = region, group = group), color = "white", fill = "pink") +
  guides(fill = FALSE) +
  geom_point(data = locMappings, aes(x = long, y = lat), color = "purple", size = 3) +
```

```
geom_text(data = locMappings, aes(x = long, y = lat, label = city), check_overlap = FALSE, angle = 0,
```



```
# overall ratings in US Apple
rev %>%
  filter(country == "USA") %>%
  filter(company == "apple") %>%
  group_by(city) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = "city") %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  ggplot(aes(x = reorder(city, mean_overall_rating_city), y = mean_overall_rating_city, color = city)) +
  geom_point(size = 5) +
  coord_flip() +
  labs(title = "Apple employee reviews by US cities", x = "Overall Ratings", y = "US cities", color = "city")
```

Apple employee reviews by US cities



```
# map for apple
county <- map_data("county")
states <- map_data("state")

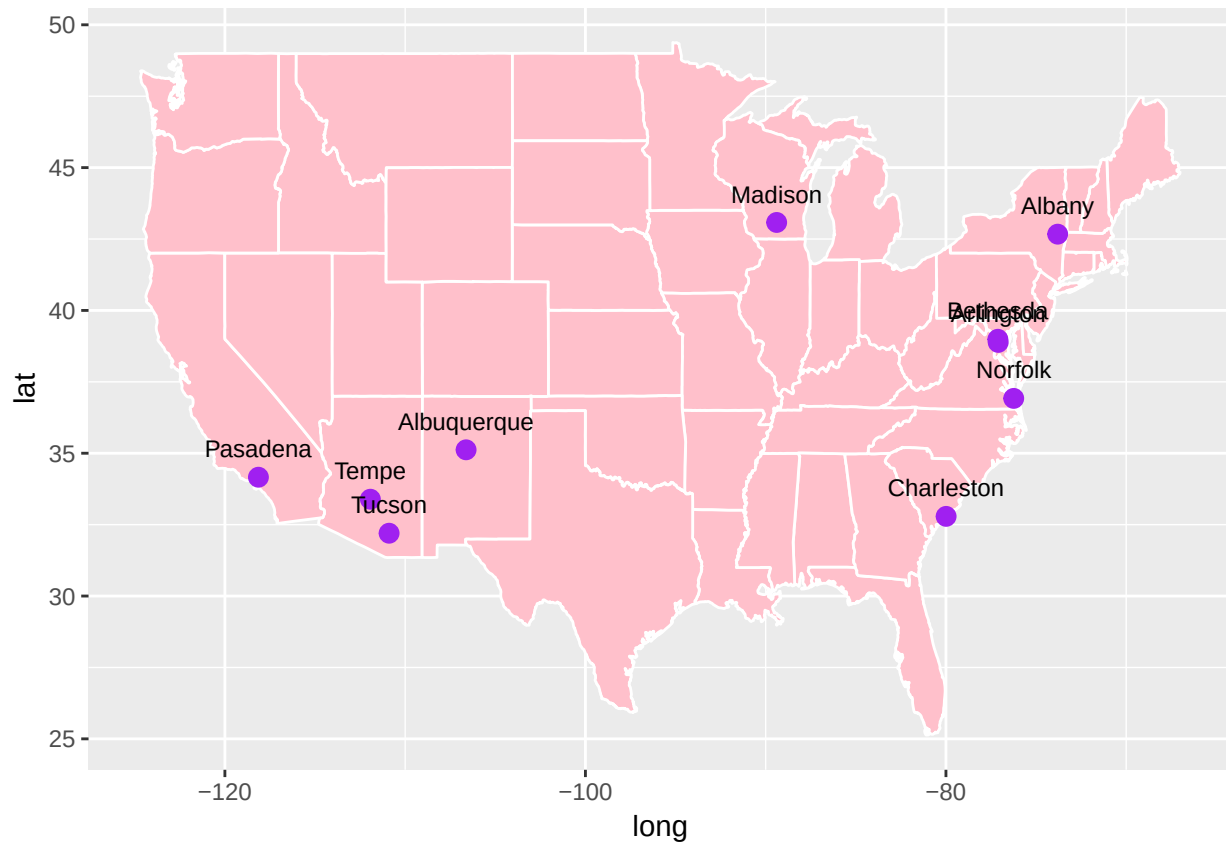
loc <- us.cities %>%
  separate(name, c("city", "state"), sep = " ") %>%
  filter(long >= -130)
```

```
## Warning: Expected 2 pieces. Additional pieces discarded in 311 rows [11,
## 16, 24, 28, 29, 33, 39, 42, 54, 56, 58, 59, 66, 67, 88, 92, 93, 96, 97,
## 103, ...].
```

```
locMappings <- inner_join(rev, loc, by = c("city", "state")) %>%
  filter(country == "USA") %>%
  filter(company == "apple") %>%
  group_by(city, state) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = c("city", "state")) %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  left_join(loc, by = c("city", "state"))
```

```
ggplot(states) +
  geom_polygon(aes(x = long, y = lat, fill = region, group = group), color = "white", fill = "pink") +
  guides(fill = FALSE) +
  geom_point(data = locMappings, aes(x = long, y = lat), color = "purple", size = 3) +
```

```
geom_text(data = locMappings, aes(x = long, y = lat, label = city), check_overlap = FALSE, angle = 0,
```



```
# overall ratings in US Microsoft
rev %>%
  filter(country == "USA") %>%
  filter(company == "microsoft") %>%
  group_by(city) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = "city") %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  ggplot(aes(x = reorder(city, mean_overall_rating_city), y = mean_overall_rating_city, color = city)) +
  geom_point(size = 5) +
  coord_flip() +
  labs(title = "Microsoft employee reviews by US cities", x = "Overall Ratings", y = "US cities", color
```



```
# map for apple
county <- map_data("county")
states <- map_data("state")

loc <- us.cities %>%
  separate(name, c("city", "state"), sep = " ") %>%
  filter(long >= -130)

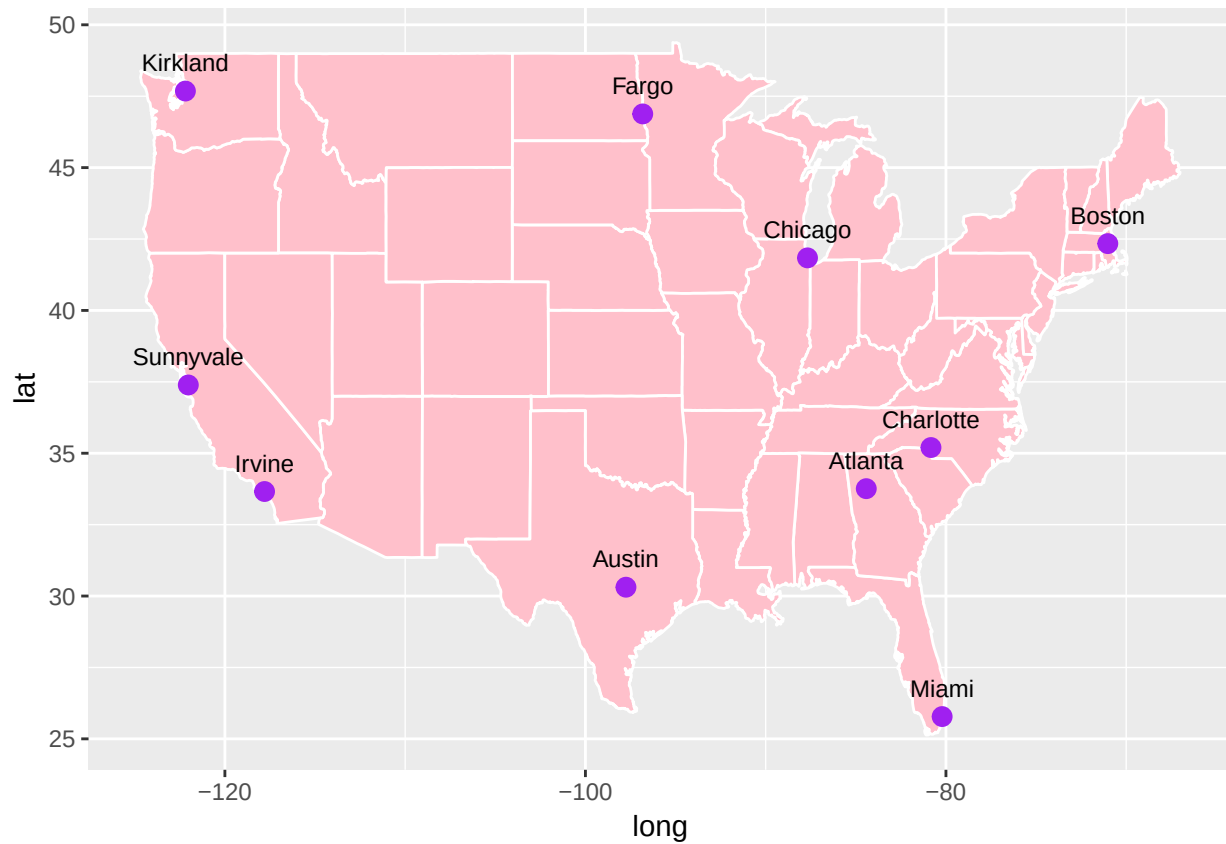
## Warning: Expected 2 pieces. Additional pieces discarded in 311 rows [11,
## 16, 24, 28, 29, 33, 39, 42, 54, 56, 58, 59, 66, 67, 88, 92, 93, 96, 97,
## 103, ...].

locMappings <- inner_join(rev, loc, by = c("city", "state")) %>%
  filter(country == "USA") %>%
  filter(company == "microsoft") %>%
  group_by(city, state) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = c("city", "state")) %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  left_join(loc, by = c("city", "state"))

ggplot(states) +
  geom_polygon(aes(x = long, y = lat, fill = region, group = group), color = "white", fill = "pink") +
  guides(fill = FALSE) +
  geom_point(data = locMappings, aes(x = long, y = lat), color = "purple", size = 3) +
```

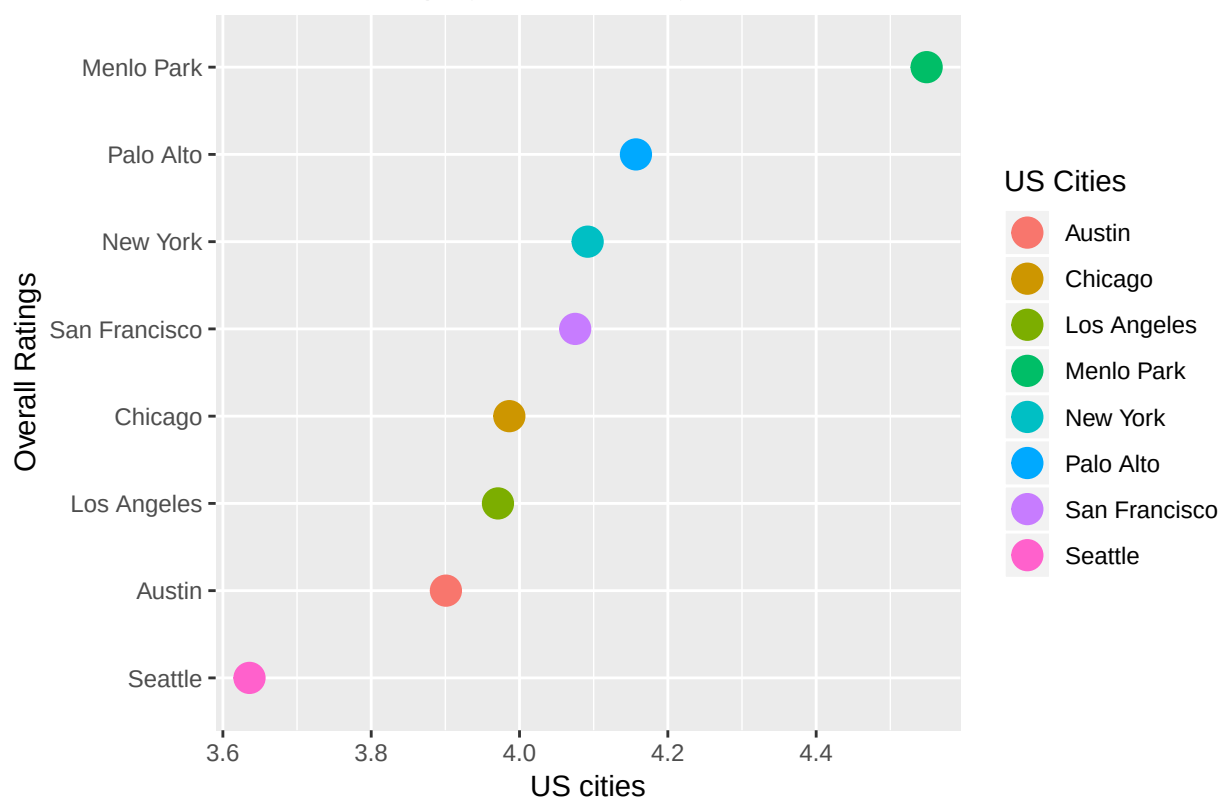


```
geom_text(data = locMappings, aes(x = long, y = lat, label = city), check_overlap = FALSE, angle = 0,
```



```
# overall ratings in US Facebook
rev %>%
  filter(country == "USA") %>%
  filter(company == "facebook") %>%
  group_by(city) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = "city") %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  ggplot(aes(x = reorder(city, mean_overall_rating_city), y = mean_overall_rating_city, color = city)) +
  geom_point(size = 5) +
  coord_flip() +
  labs(title = "Facebook employee reviews by US cities", x = "Overall Ratings", y = "US cities", color = "city")
```

Facebook employee reviews by US cities



```
# map for facebook
county <- map_data("county")
states <- map_data("state")

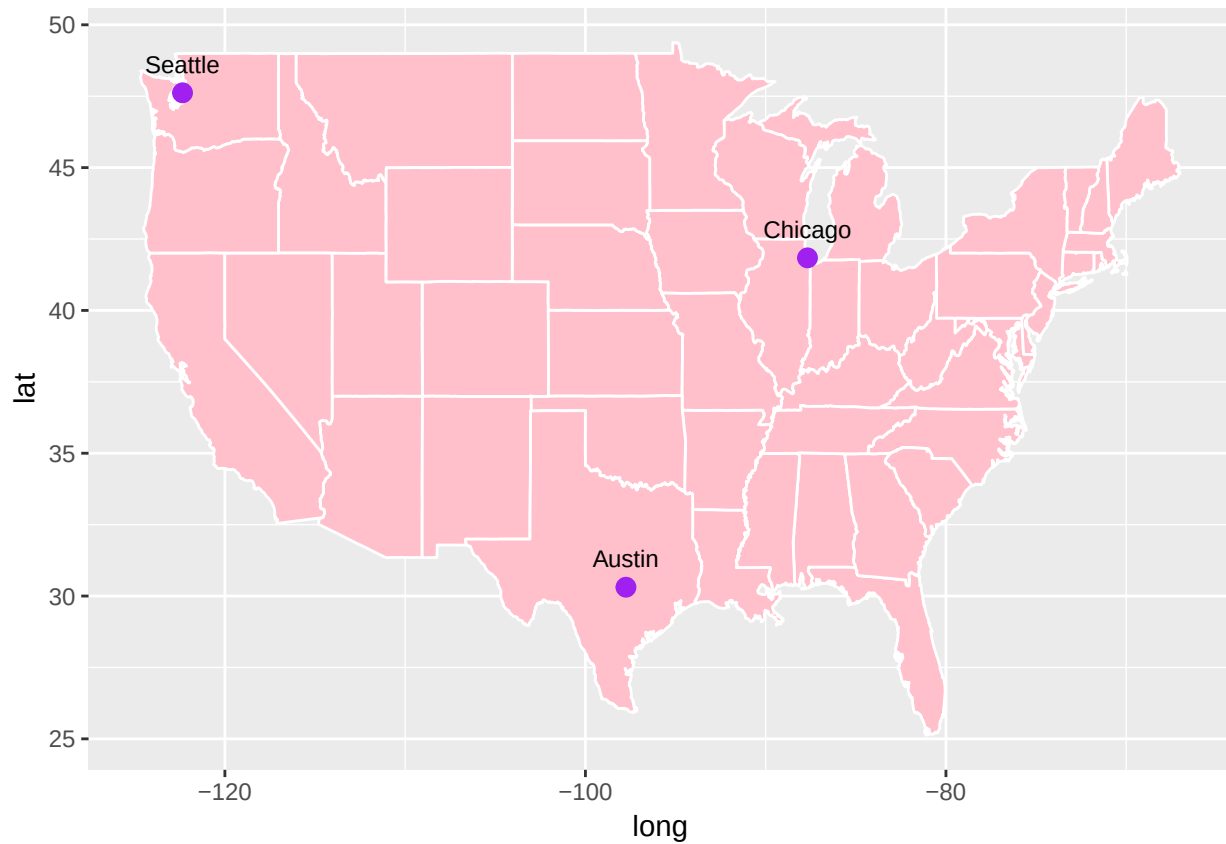
loc <- us.cities %>%
  separate(name, c("city", "state"), sep = " ") %>%
  filter(long >= -130)

## Warning: Expected 2 pieces. Additional pieces discarded in 311 rows [11,
## 16, 24, 28, 29, 33, 39, 42, 54, 56, 58, 59, 66, 67, 88, 92, 93, 96, 97,
## 103, ...].

locMappings <- inner_join(rev, loc, by = c("city", "state")) %>%
  filter(country == "USA") %>%
  filter(company == "facebook") %>%
  group_by(city, state) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = c("city", "state")) %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  left_join(loc, by = c("city", "state"))

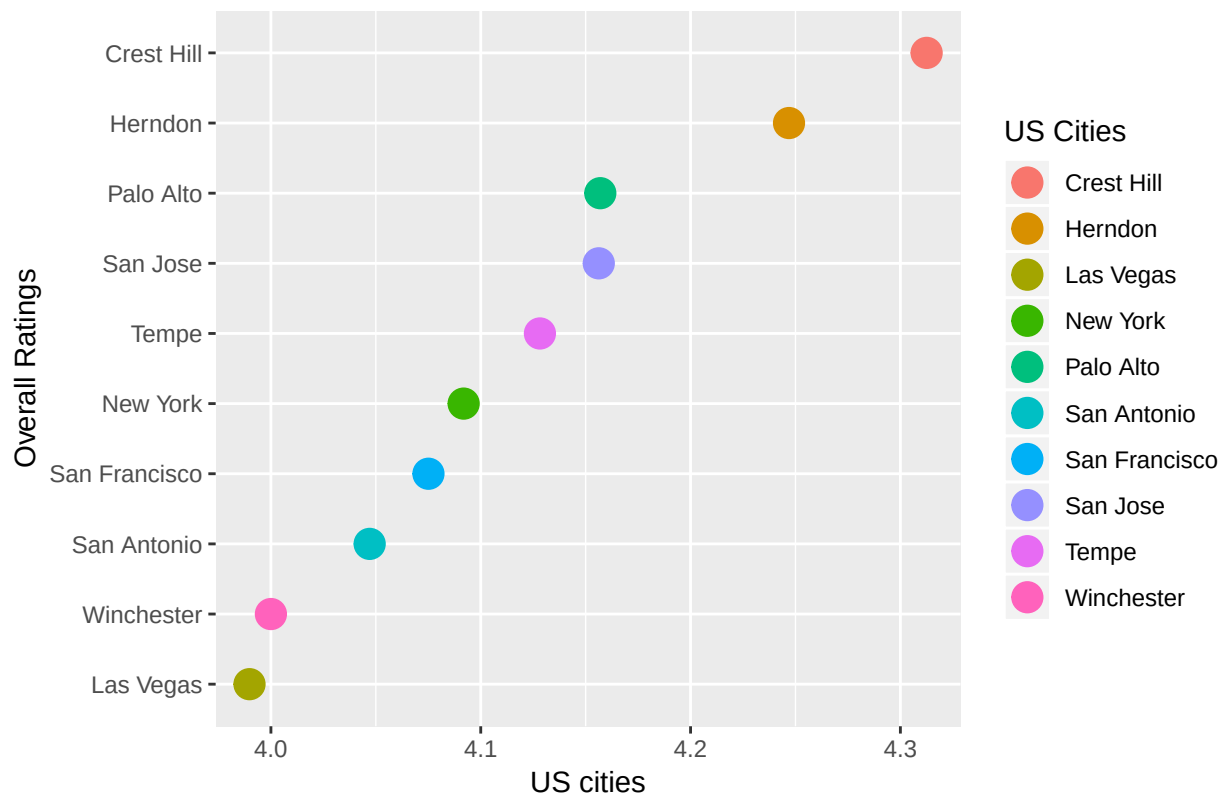
ggplot(states) +
  geom_polygon(aes(x = long, y = lat, fill = region, group = group), color = "white", fill = "pink") +
  guides(fill = FALSE) +
  geom_point(data = locMappings, aes(x = long, y = lat), color = "purple", size = 3) +
```

```
geom_text(data = locMappings, aes(x = long, y = lat, label = city), check_overlap = FALSE, angle = 0,
```



```
# overall ratings in US Amazon
rev %>%
  filter(country == "USA") %>%
  filter(company == "amazon") %>%
  group_by(city) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = "city") %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  ggplot(aes(x = reorder(city, mean_overall_rating_city), y = mean_overall_rating_city, color = city)) +
  geom_point(size = 5) +
  coord_flip() +
  labs(title = "Amazon employee reviews by US cities", x = "Overall Ratings", y = "US cities", color = "city")
```

Amazon employee reviews by US cities



```
# map for facebook
county <- map_data("county")
states <- map_data("state")

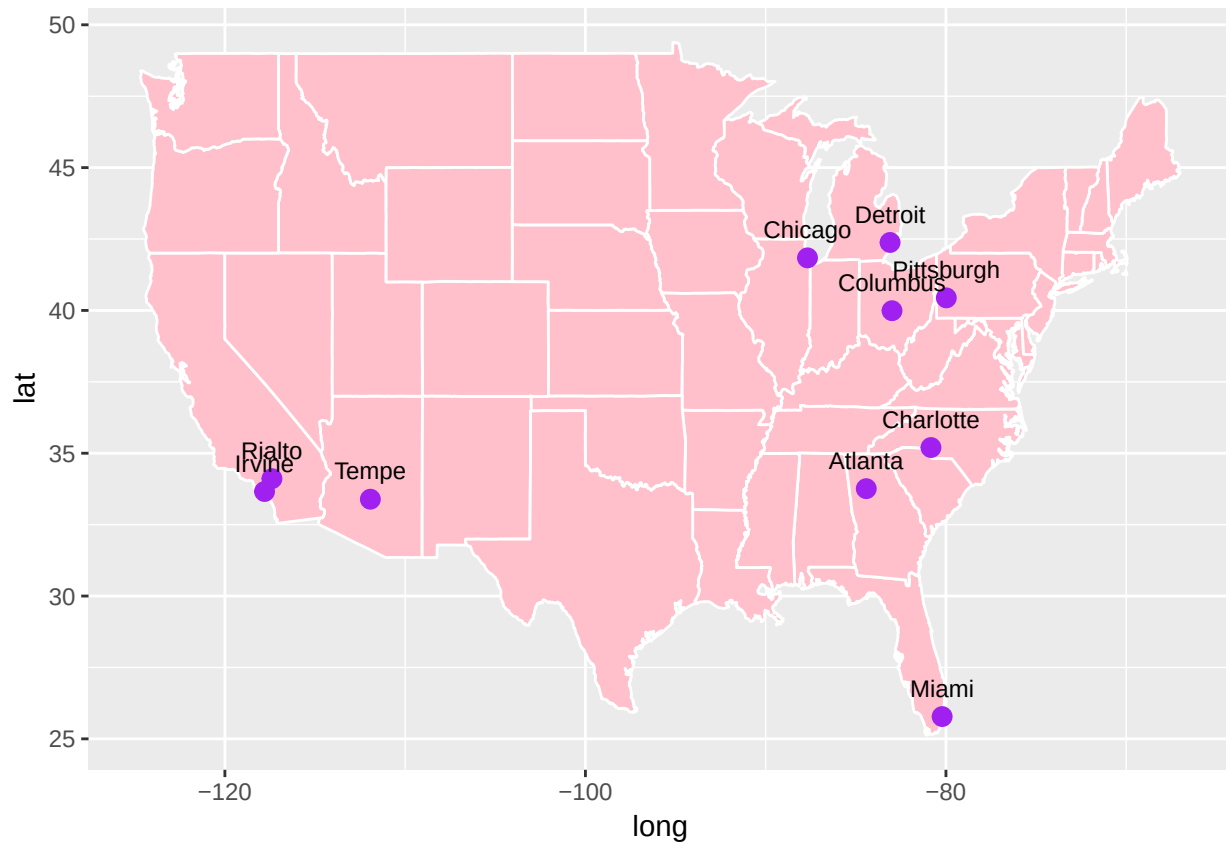
loc <- us.cities %>%
  separate(name, c("city", "state"), sep = " ") %>%
  filter(long >= -130)
```

```
## Warning: Expected 2 pieces. Additional pieces discarded in 311 rows [11,
## 16, 24, 28, 29, 33, 39, 42, 54, 56, 58, 59, 66, 67, 88, 92, 93, 96, 97,
## 103, ...].
```

```
locMappings <- inner_join(rev, loc, by = c("city", "state")) %>%
  filter(country == "USA") %>%
  filter(company == "amazon") %>%
  group_by(city, state) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = c("city", "state")) %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  left_join(loc, by = c("city", "state"))
```

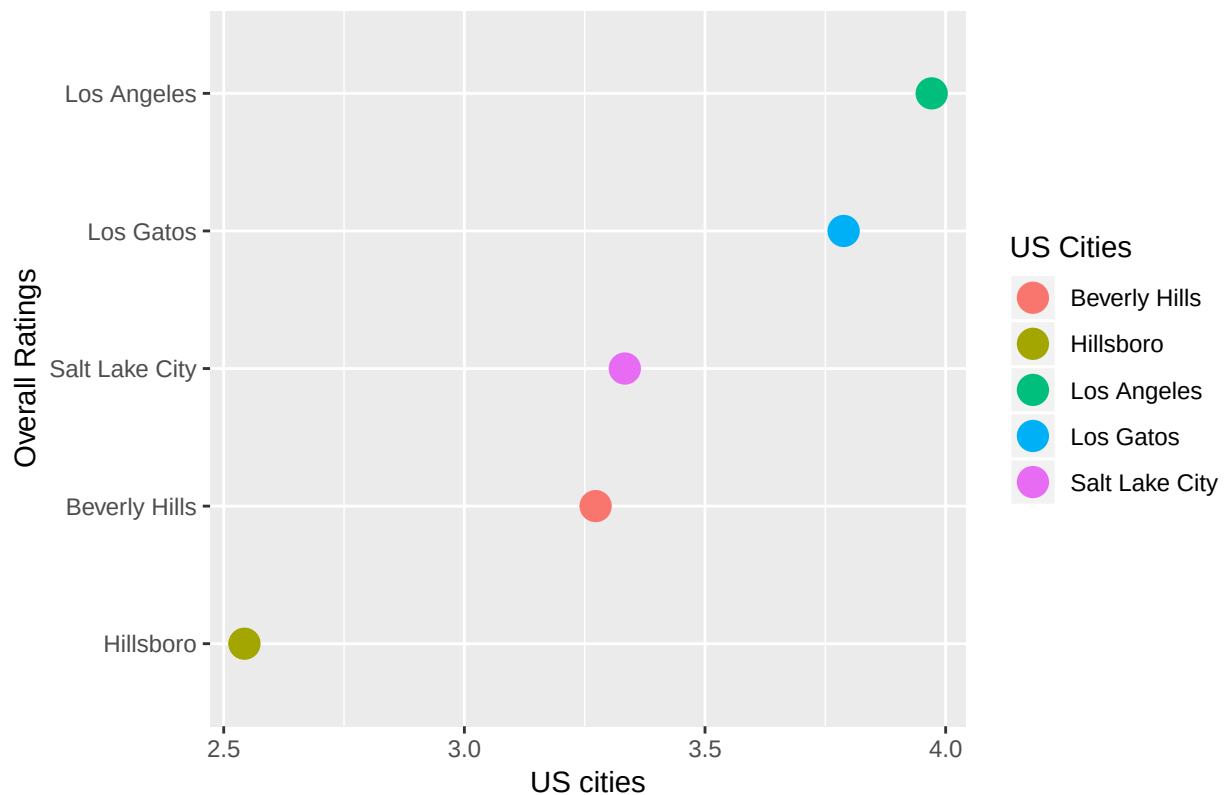
```
ggplot(states) +
  geom_polygon(aes(x = long, y = lat, fill = region, group = group), color = "white", fill = "pink") +
  guides(fill = FALSE) +
  geom_point(data = locMappings, aes(x = long, y = lat), color = "purple", size = 3) +
```

```
geom_text(data = locMappings, aes(x = long, y = lat, label = city), check_overlap = FALSE, angle = 0,
```



```
# overall ratings in US Netflix
rev %>%
  filter(country == "USA") %>%
  filter(company == "netflix") %>%
  group_by(city) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = "city") %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  ggplot(aes(x = reorder(city, mean_overall_rating_city), y = mean_overall_rating_city, color = city)) +
  geom_point(size = 5) +
  coord_flip() +
  labs(title = "Netflix employee reviews by US cities", x = "Overall Ratings", y = "US cities", color =
```

Netflix employee reviews by US cities



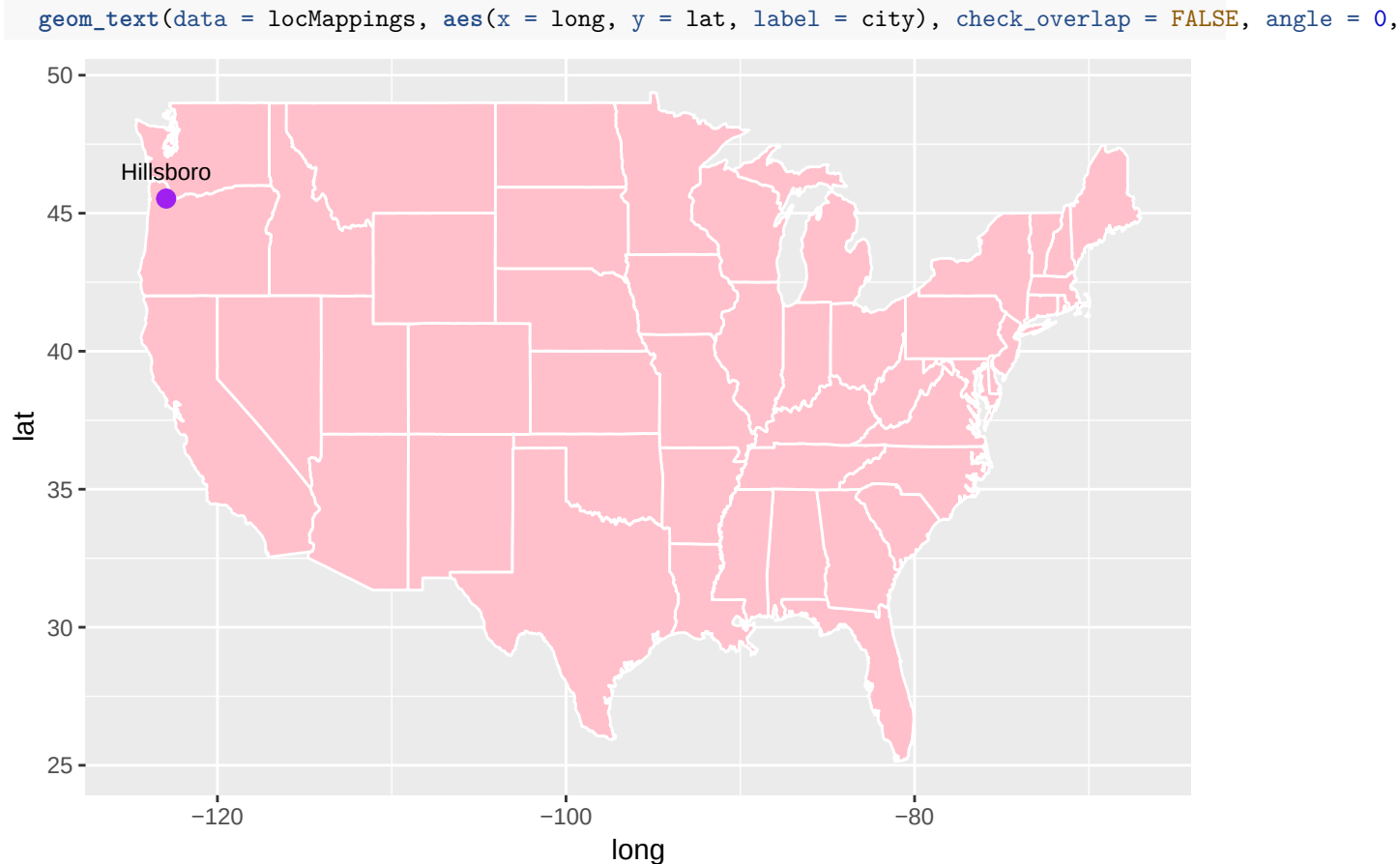
```
# map for facebook
county <- map_data("county")
states <- map_data("state")

loc <- us.cities %>%
  separate(name, c("city", "state"), sep = " ") %>%
  filter(long >= -130)

## Warning: Expected 2 pieces. Additional pieces discarded in 311 rows [11,
## 16, 24, 28, 29, 33, 39, 42, 54, 56, 58, 59, 66, 67, 88, 92, 93, 96, 97,
## 103, ...].

locMappings <- inner_join(rev, loc, by = c("city", "state")) %>%
  filter(country == "USA") %>%
  filter(company == "netflix") %>%
  group_by(city, state) %>%
  count() %>%
  filter(n > 10) %>%
  left_join(rev, by = c("city", "state")) %>%
  summarise(mean_overall_rating_city = mean(overall_ratings)) %>%
  arrange(desc(mean_overall_rating_city)) %>%
  head(10) %>%
  left_join(loc, by = c("city", "state"))

ggplot(states) +
  geom_polygon(aes(x = long, y = lat, fill = region, group = group), color = "white", fill = "pink") +
  guides(fill = FALSE) +
  geom_point(data = locMappings, aes(x = long, y = lat), color = "purple", size = 3) +
```



Job Position and Level Reviews

```
rev_rating_job <- rev %>%
  mutate(Engineer = str_detect(title, "Engineer") | str_detect(title, "SDET") | str_detect(title, "SDE1") |
    str_detect(title, "SDE2") | str_detect(title, "Software Developer") | str_detect(title, "Software Engineer"),
    Manager = str_detect(title, "Manager"),
    Intern = str_detect(title, "Intern"),
    Analyst = str_detect(title, "Analyst"),
    Executive = str_detect(title, "Executive"),
    Customer_service = str_detect(title, "CSA"))

rev_rating_job <- rev_rating_job %>%
  unite(job_position, c(Engineer, Manager, Intern, Analyst, Executive, Customer_service)) %>%
  mutate(job_position = str_replace(job_position, "TRUE_FALSE_FALSE_FALSE_FALSE_FALSE", "Software Engineer"),
    job_position = str_replace(job_position, "FALSE_TRUE_FALSE_FALSE_FALSE_FALSE", "Manager"),
    job_position = str_replace(job_position, "TRUE_TRUE_FALSE_FALSE_FALSE_FALSE", "Manager"),
    job_position = str_replace(job_position, "FALSE_TRUE_FALSE_FALSE_TRUE_FALSE", "Manager"),
    job_position = str_replace(job_position, "FALSE_FALSE_TRUE_FALSE_FALSE_FALSE", "Intern"),
    job_position = str_replace(job_position, "TRUE_FALSE_TRUE_FALSE_FALSE_FALSE", "Intern"),
    job_position = str_replace(job_position, "FALSE_FALSE_TRUE_TRUE_FALSE_FALSE", "Intern"),
    job_position = str_replace(job_position, "FALSE_FALSE_TRUE_FALSE_TRUE_FALSE", "Intern"),
    job_position = str_replace(job_position, "FALSE_TRUE_TRUE_FALSE_FALSE_FALSE", "Intern"),
```

```

job_position = str_replace(job_position, "TRUE_TRUE_TRUE_FALSE_FALSE_FALSE", "Intern"),
job_position = str_replace(job_position, "FALSE_FALSE_FALSE_TRUE_FALSE_FALSE", "Analyst"),
job_position = str_replace(job_position, "FALSE_FALSE_FALSE_TRUE_TRUE_FALSE", "Analyst"),
job_position = str_replace(job_position, "FALSE_TRUE_FALSE_TRUE_FALSE_FALSE", "Analyst"),
job_position = str_replace(job_position, "TRUE_FALSE_FALSE_TRUE_FALSE_FALSE", "Analyst"),
job_position = str_replace(job_position, "FALSE_FALSE_FALSE_FALSE_TRUE_FALSE", "Executive Board"),
job_position = str_replace(job_position, "FALSE_FALSE_FALSE_FALSE_FALSE_TRUE", "Customer Service"),
job_position = str_replace(job_position, "FALSE_FALSE_FALSE_FALSE_FALSE_FALSE", "Other") )

rev_rating_job <- rev_rating_job %>%
  group_by(job_position) %>%
  summarise(overall_ratings = mean(overall_ratings, na.rm = TRUE),
            work_balance_stars = mean(work_balance_stars, na.rm = TRUE),
            culture_values_stars = mean(culture_values_stars, na.rm = TRUE),
            career_opportunities_stars = mean(carrer_opportunities_stars, na.rm = TRUE),
            comp_benefit_stars = mean(comp_benefit_stars, na.rm = TRUE),
            senior_mangemnet_stars = mean(senior_mangemnet_stars, na.rm = TRUE))

ggplot(rev_rating_job, aes(x = reorder(job_position, overall_ratings), y = overall_ratings , col = job_position))
  geom_count(size=7) +
  coord_flip() +
  labs(title="Overall Ratings vs Job Position", x = "Overall Ratings", y = "Job Position")

```



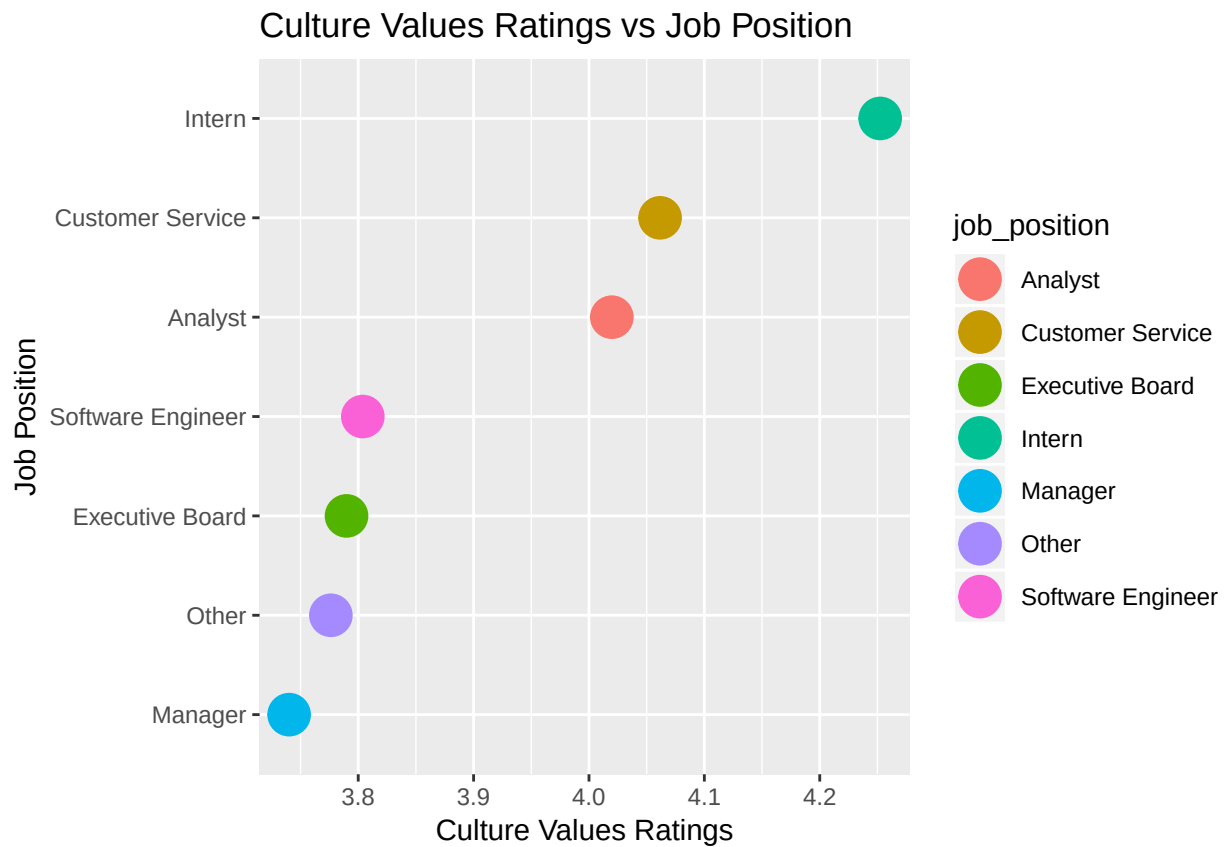
```

ggplot(rev_rating_job, aes(x = reorder(job_position, work_balance_stars), y = work_balance_stars, col = job_position))
  geom_count(size=7) +
  coord_flip() +
  labs(title="Work Balance Ratings vs Job Position", y = "Work Balance Ratings", x = "Job Position")

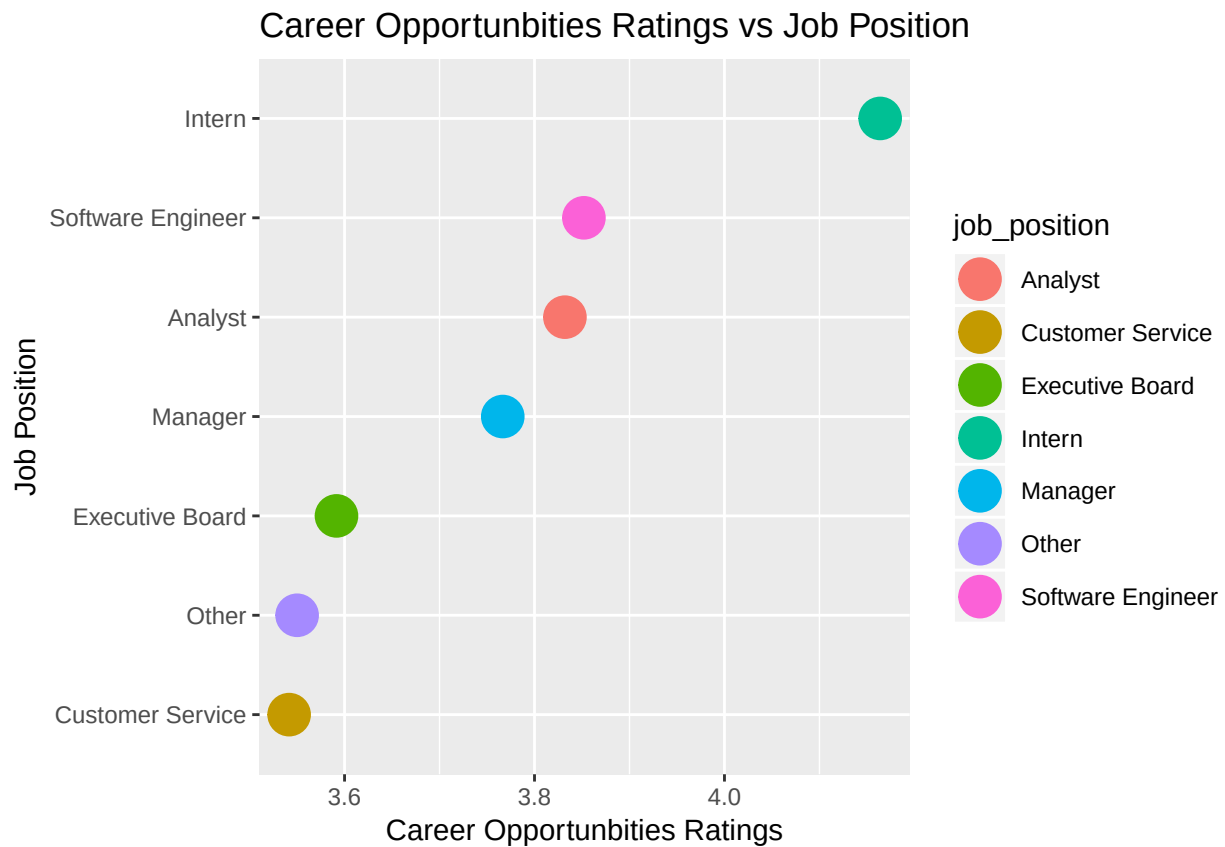
```




```
ggplot(rev_rating_job, aes(x = reorder(job_position, culture_values_stars), y = culture_values_stars , 
  geom_count(size=7) + 
  coord_flip() + 
  labs(title="Culture Values Ratings vs Job Position", y = "Culture Values Ratings", x = "Job Position")
```



```
ggplot(rev_rating_job, aes(x = reorder(job_position, career_opportunities_stars), y = career_opportunities_stars)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Career Opportunbities Ratings vs Job Position", y = "Career Opportunbities Ratings", x = "Job Position")
```



```
ggplot(rev_rating_job, aes(x = reorder(job_position, comp_benefit_stars), y = comp_benefit_stars , col = job_position)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Company Benefit Ratings vs Job Position", y = "Company Benefit Ratings", x = "Job Position")
```



```
ggplot(rev_rating_job, aes(x = reorder(job_position, senior_mangemnet_stars), y = senior_mangemnet_stars)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Senior Management Ratings vs Job Position", y = "Senior Management Ratings", x = "Job Position")
```



```

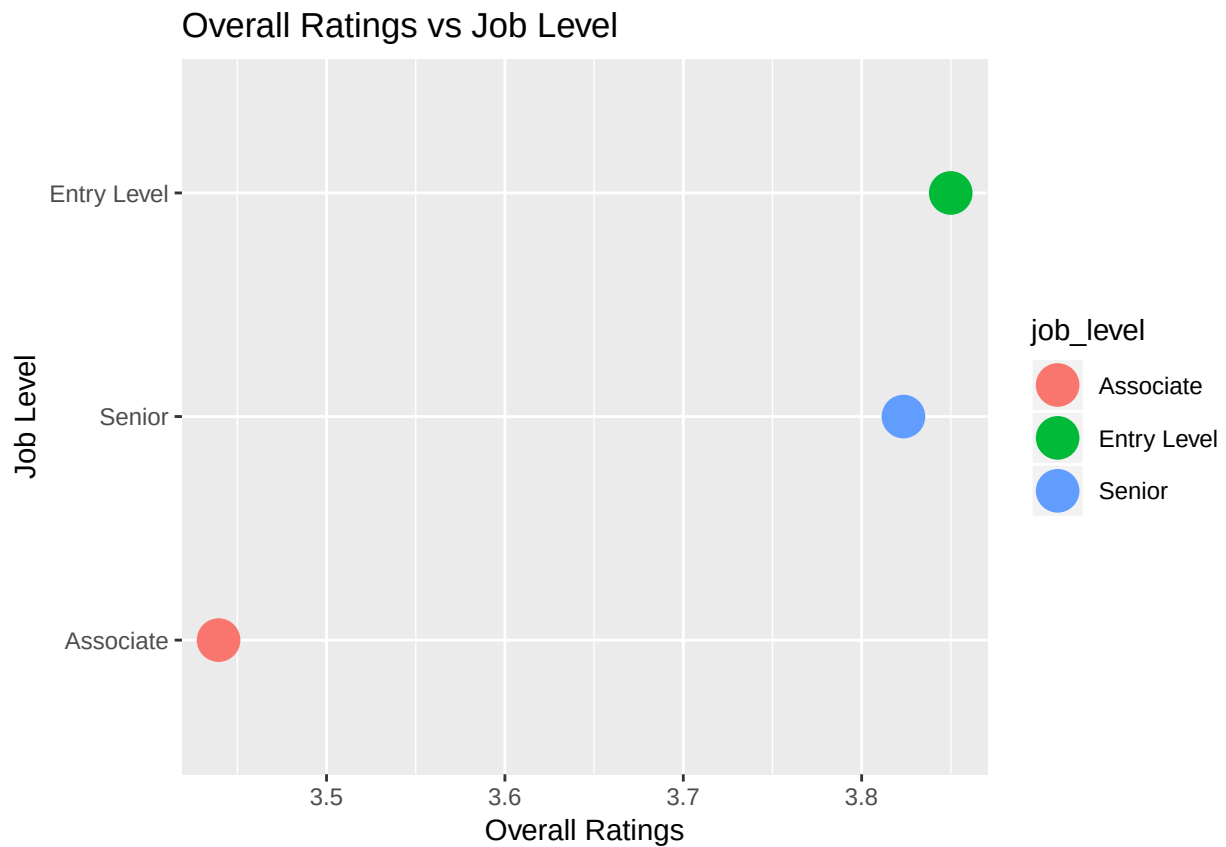
rev_rating_jobLevel <- rev %>%
  mutate(Senior = str_detect(title, "[Ss]enior"),
         Associate = str_detect(title, "[Aa]ssociate"),
         Entry_level = !(str_detect(title, "[Ss]enior") | str_detect(title, "[Aa]ssociate"))) )

rev_rating_jobLevel <- rev_rating_jobLevel %>%
  unite(job_level, c(Senior, Associate, Entry_level)) %>%
  mutate(job_level = str_replace(job_level, "TRUE_FALSE_FALSE", "Senior"),
         job_level = str_replace(job_level, "TRUE_TRUE_FALSE", "Senior"),
         job_level = str_replace(job_level, "FALSE_TRUE_FALSE", "Associate"),
         job_level = str_replace(job_level, "FALSE_FALSE_TRUE", "Entry Level") )

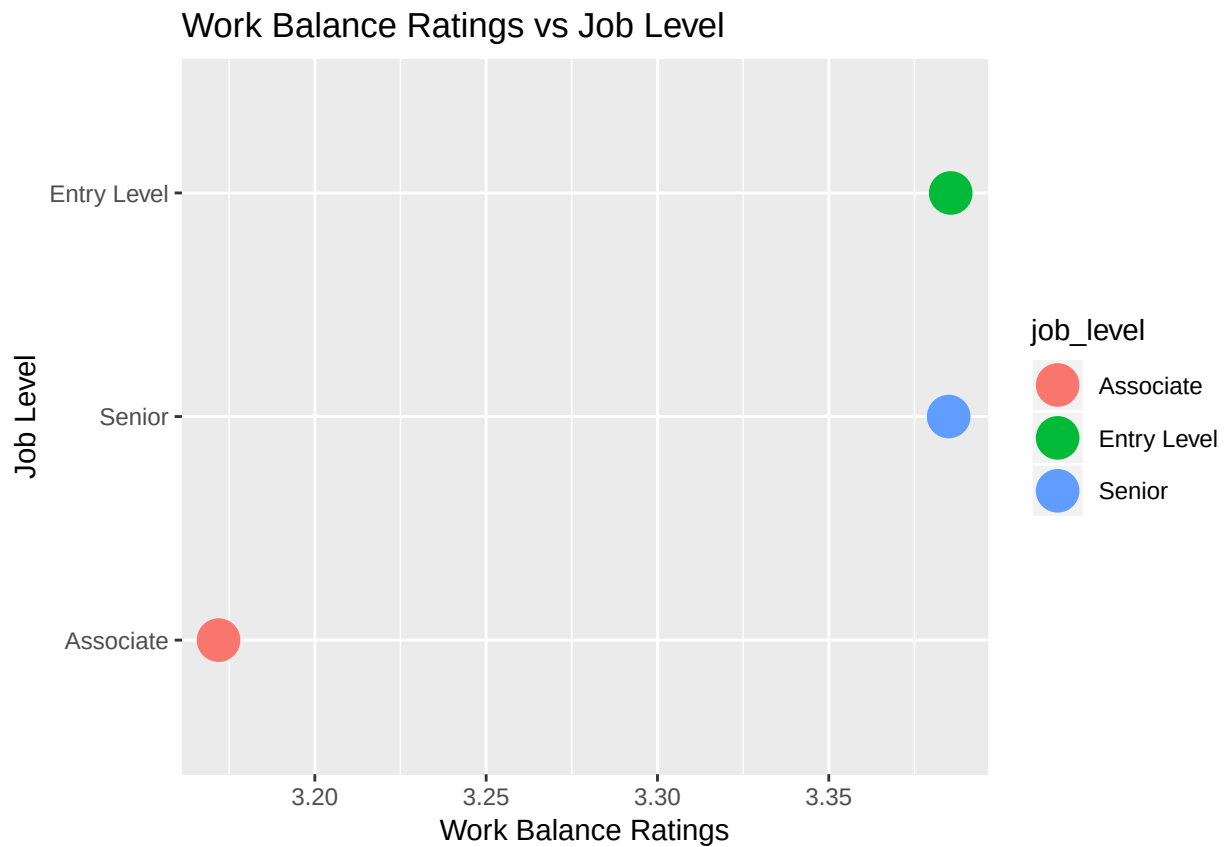
rev_rating_jobLevel <- rev_rating_jobLevel %>%
  group_by(job_level) %>%
  summarise(overall_ratings = mean(overall_ratings, na.rm = TRUE),
            work_balance_stars = mean(work_balance_stars, na.rm = TRUE),
            culture_values_stars = mean(culture_values_stars, na.rm = TRUE),
            career_opportunities_stars = mean(carrer_opportunities_stars, na.rm = TRUE),
            comp_benefit_stars = mean(comp_benefit_stars, na.rm = TRUE),
            senior_mangemnet_stars = mean(senior_mangemnet_stars, na.rm = TRUE))

ggplot(rev_rating_jobLevel, aes(x = reorder(job_level, overall_ratings), y = overall_ratings , col = job_level)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Overall Ratings vs Job Level", y = "Overall Ratings", x = "Job Level")

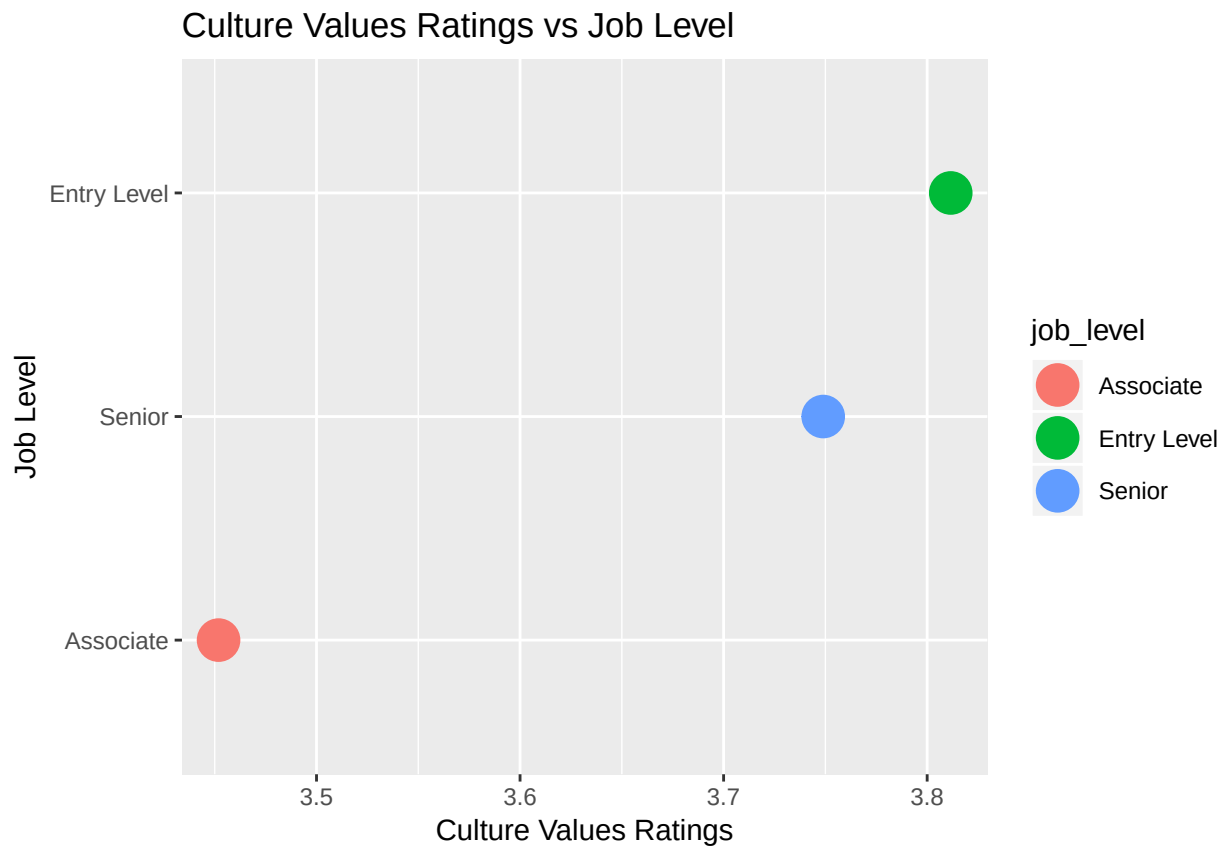
```



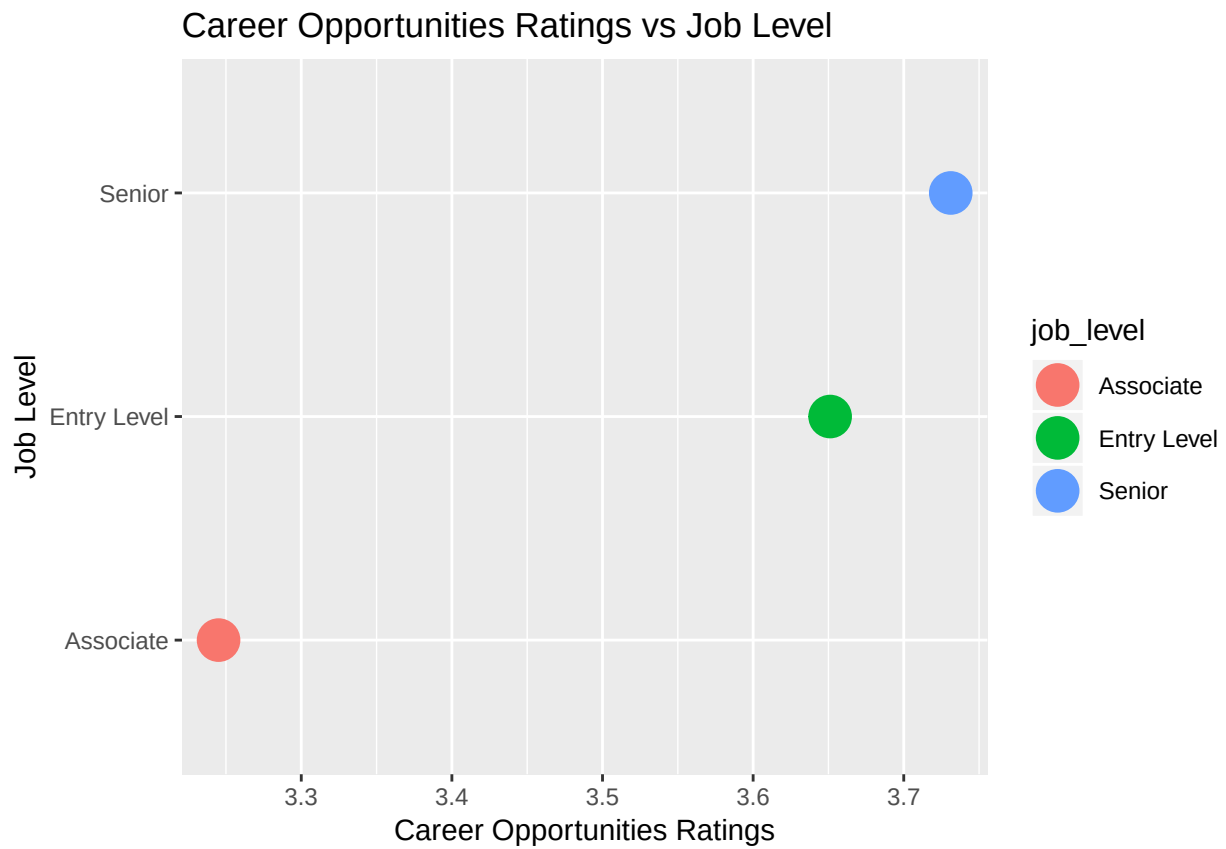
```
ggplot(rev_rating_jobLevel, aes(x = reorder(job_level, work_balance_stars), y = work_balance_stars, col = job_level)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Work Balance Ratings vs Job Level", y = "Work Balance Ratings", x = "Job Level")
```



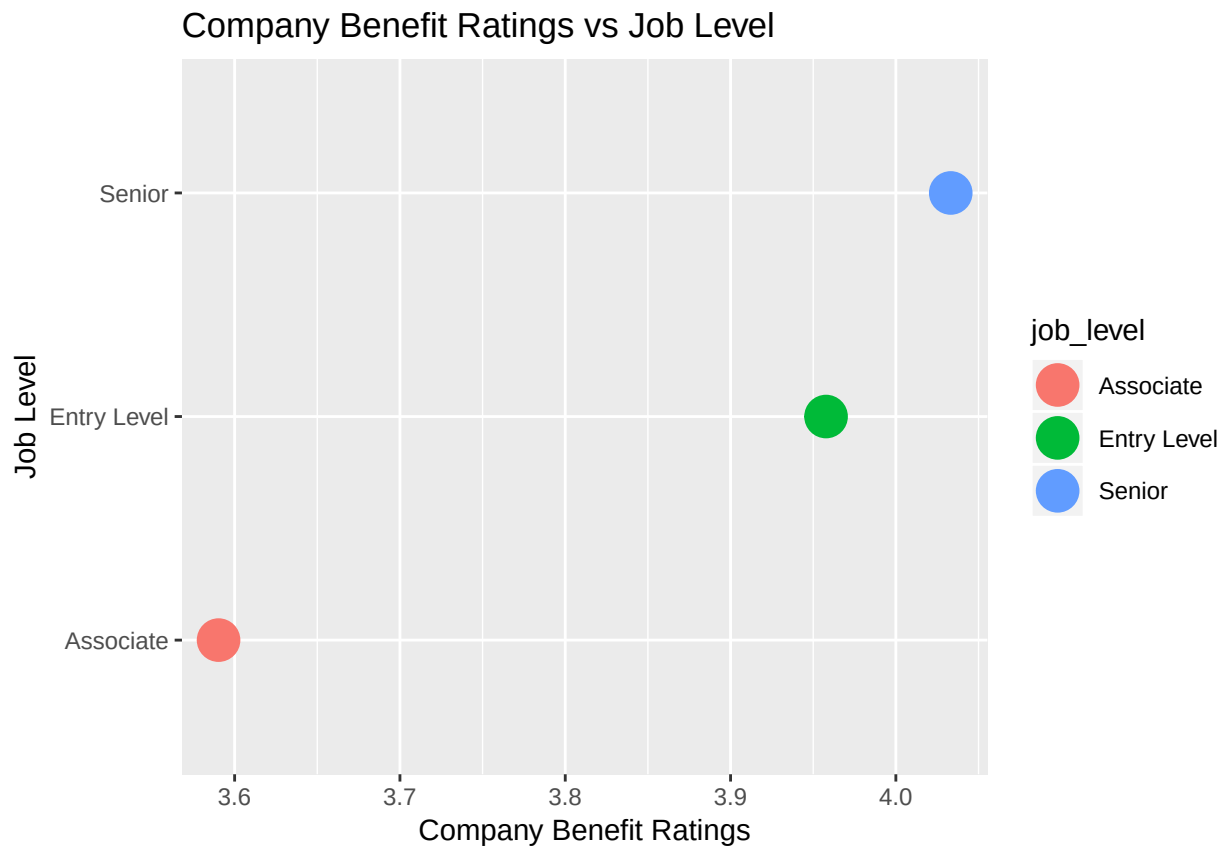
```
ggplot(rev_rating_jobLevel, aes(x = reorder(job_level, culture_values_stars), y = culture_values_stars)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Culture Values Ratings vs Job Level", y = "Culture Values Ratings", x = "Job Level")
```



```
ggplot(rev_rating_jobLevel, aes(x = reorder(job_level, career_opportunities_stars), y = career_opportunities_stars)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Career Opportunities Ratings vs Job Level", y = "Career Opportunities Ratings", x = "Job Level")
```

```
ggplot(rev_rating_jobLevel, aes(x = reorder(job_level, comp_benefit_stars), y = comp_benefit_stars , color = job_level)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Company Benefit Ratings vs Job Level", y = "Company Benefit Ratings", x = "Job Level")
```



```
ggplot(rev_rating_jobLevel, aes(x = reorder(job_level, senior_mangemnet_stars), y = senior_mangemnet_stars)) +
  geom_count(size=7) +
  coord_flip() +
  labs(title="Senior Management Ratings vs Job Level", y = "Senior Management Ratings", x = "Job Level")
```

