

تحلیل و بررسی دنیای پیتزا

امتیازها بازخوردها و روندها

بر اساس داده های سایت TidyTuesday



فهرست محتوا

1

شروع سفر با داده ها

معرفی دیتاست پیتزا و آماده سازی برای تحلیل

3

راز های نهفته در نمودار ها

کشف الگو و روند ها با استفاده از نمودار های پیچیده

2

نقاشی با اعداد

رنگ آمیزی داستان داده ها با نمودار های ساده

4

نتیجه گیری خوشمزه

جمع بندی داستان داده ها پایین هر نمودار



معرفی دیتاست پیتزا

به دنیای پیتزا های خوشمزه نیویورک خوش اومدی!
دنیایی که هر لقمه و هر امتیاز و هر عدد داستانی برای گفتن دارد
در این سفر لذیذ ما از خیابان های پر هیاهوی منتهن تا گوشه های دنج بروکلین رد
پاهای پیتزا هایی را دنبال می کنیم که یا محبوب دل ها هستند یا قربانی نقد های تند
و در این میان اطلاعاتی مثل امتیاز های چندگانه (مردم و منتقدین و کارکنان) /
قیمت هر پیتزا / نوع پیتزا / مکان دقیق رستوران ها ما را همراهی میکند .
و هدف ما رمزگشایی از رابطه میان کیفیت و قیمت و محبوبیت پیتزا ها با چاشنی از
نمودار ها و ادویه ای از تحلیل و عطری از اشتیاق برای کشف روابط است

نحوه دسترسی خوشمزه :

فراخوانی داده های پیتزا برای تحلیل در R

```
pizza_jared <-  
readr::read_csv("https://raw.githubusercontent.com/  
fordata-science/tidy-tuesday/main/data/2019/2019-10-  
01/pizza_jared.csv")  
pizza_barstool <-  
readr::read_csv("https://raw.githubusercontent.com/  
fordata-science/tidy-tuesday/main/data/2019/2019-10-  
01/pizza_barstool.csv")  
pizza_datafiniti <-  
readr::read_csv("https://raw.githubusercontent.com/  
fordata-science/tidy-tuesday/main/data/2019/2019-10-  
01/pizza_datafiniti.csv")
```

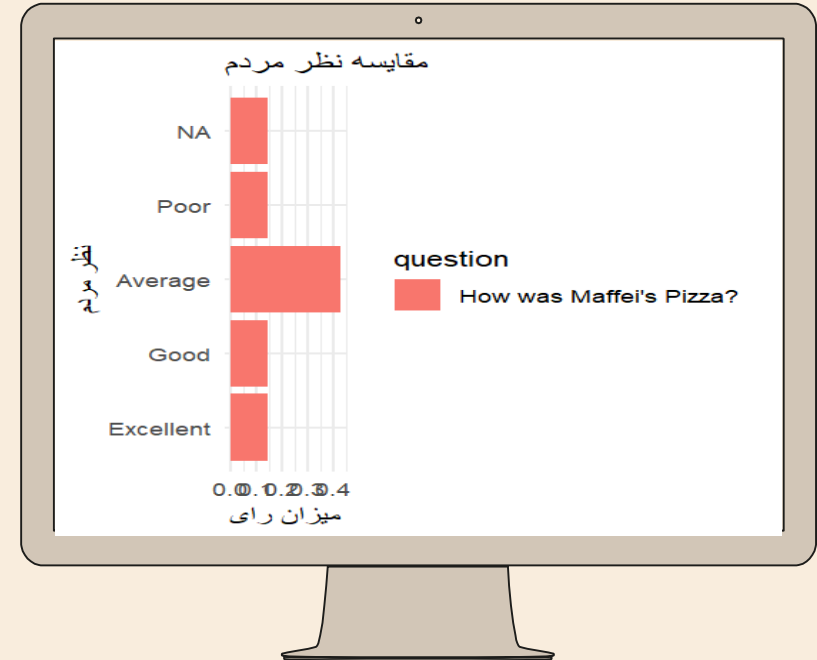


داده های ما شامل سه مجموعه هستند که شامل نظرسنجی های مردمی راجب پیتزا و اطلاعات تجاری رستوران های پیتزا و همچنین شامل نقد و امتیازدهی یک منتقد معروف پیتزا هست.



نتایج نظر سنجی راجب پیتزای مافی : عالی بوده یا قابل تحمل؟

```
pizza_jared %>%  
  filter(question %in% ("How was Maffei's Pizza?")) %>%  
  mutate(answer = factor(answer, levels = c("Excellent",  
    "Good", "Average", "Poor", "Never again"))) %>%  
  ggplot(aes(x = answer, y = percent, fill = question)) +  
  geom_col(position = "dodge") +  
  labs(title = "مقایسه نظر مردم",  
    x = "نظر مردم", y = "میزان رأی") +  
  theme_minimal() +  
  coord_flip ()
```



نتیجه گیری : وقتی بیشتر مردم به پیتزای مافی نمره متوسط و میدن نشون میدن که این پیتزا نه شاهکاره نه فاجعه در واقع

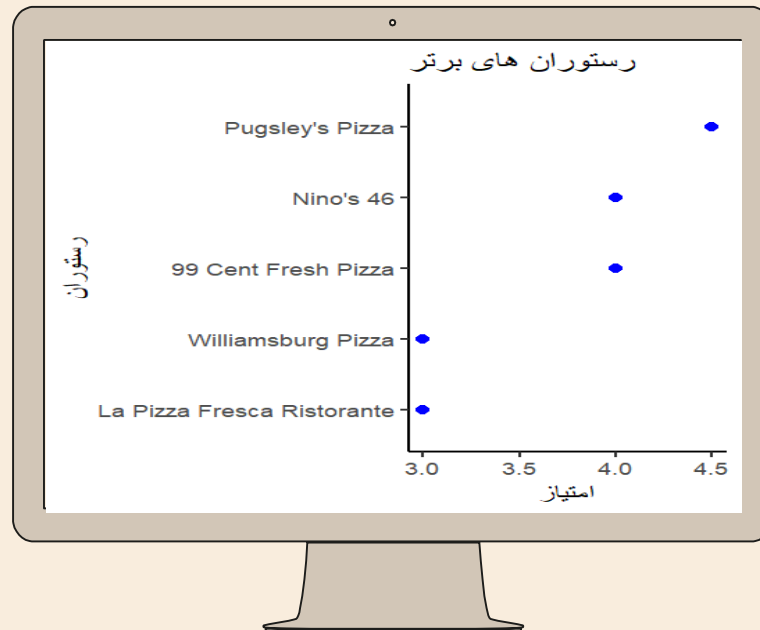
نه قراره قلب تو تسخیر کنه و نه قراره دشمن شکمت باشه فقط اومده که یه وعده رو پر کنه نه یه خاطره رو !





معرفی پنج رستوران برتر توسط منتقد معروف :

```
pizza_barstool %>%  
filter(!is.na(provider_rating)) %>%  
slice(1:5) %>%  
ggplot(aes(x = provider_rating, y = reorder(name,  
provider_rating))) +  
geom_point(color = "blue", size = 2) +  
labs(title = "رستوران های برتر"  
x = "امتیاز", y = "رستوران" +  
theme_classic()
```



نتیجه گیری : این پنج رستوران مثل سوپر قهرمان های دنیای پیتزا هستند و وقتی به منتقد حرفه ای به این پنج رستوران اینقدر نمره بالا میدهند یعنی دیگه هیچ بهونه ای برای انتخاب کردن بهترین پیتزا نیست و فقط کافیست که به امتیاز ها نگاه کنیم و با خیال راحت به پیتزای خوشمزه رو سفارش بدیم!

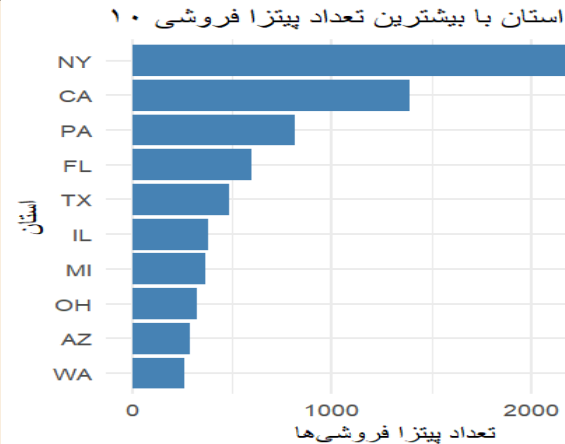


معرفی 10 استان با بیشترین تعداد پیتزا فروشی :

```

pizza_data$finiti %>%
  group_by(province) %>%
  summarise(count = n()) %>%
  arrange(desc(count)) %>%
  head(10) %>%
  ggplot(aes(x = reorder(province, count), y = count)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  coord_flip() +
  labs(title = "استان با بیشترین تعداد پیتزا فروشی ۱۰",
       x = "استان",
       y = "تعداد پیتزا فروشی‌ها") +
  theme_minimal()

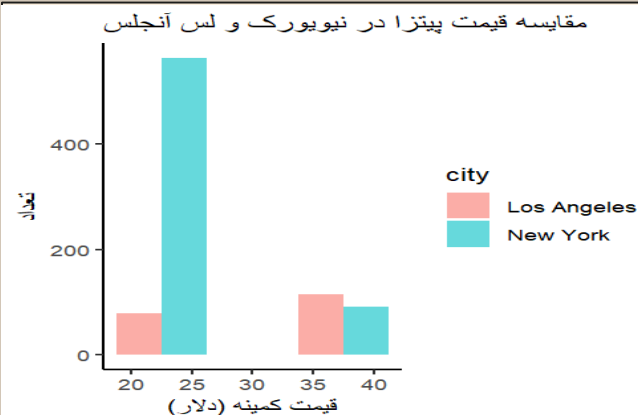
```



نتیجه گیری : این نمودار نشون میده که نیویورک نه تنها مهد فرهنگ و هنره بلکه با اختلاف پادشاه سرزمین پیتزا هاست استان های بعدی هم با این که تلاش خودشونو کردن ولی در رقابت با نیویورک کم آوردن و این نشون میده که پیتزا برای مردم نیویورک یه غذا نیست یه سبک زندگیه !

مقایسه قیمت پیتزا در نیویورک و لس آنجلس :

```
pizza_data %>%  
  filter(city %in% c("New York", "Los Angeles")) %>%  
  ggplot(aes(x = price_range_max, fill = city)) +  
  geom_histogram(position = "dodge", alpha = 0.6, bins = 30) +  
  labs(title = "مقایسه قیمت پیتزا در نیویورک و لس آنجلس",  
        x = "قیمت کمینه (دلار)",  
        y = "تعداد") +  
  theme_classic()
```

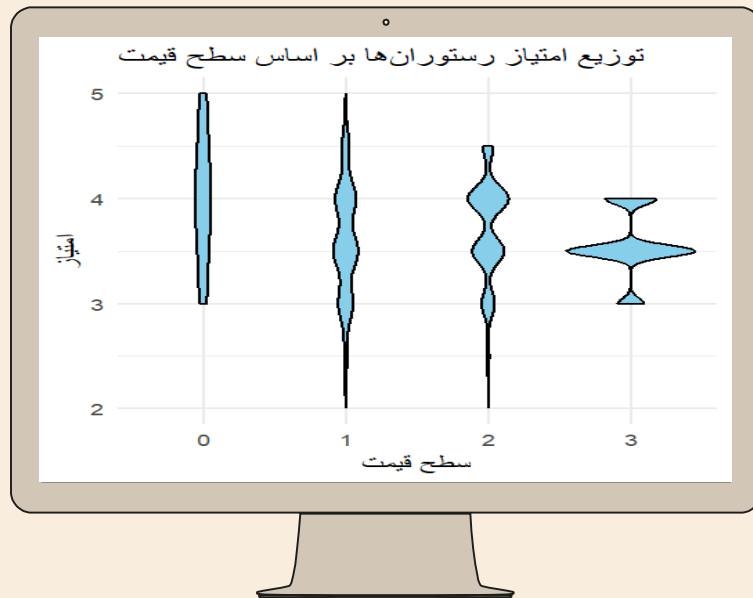


نتیجه گیری : طبق نمودار بیشترین فراوانی توی بازه 20 تا 25 دلار دیده میشه که میتونیم بگیم که توی لس آنجلس پیتزاهای ارزون تر در دسترسه و بیشترین تمرکز هم روی قیمت های 30 تا 40 دلاره که طبق مقایسه میگیریم توی نیویورک پیتزا ها گرون تره اگه دنبال یه پیتزای ارزونی که پول تو نسوزونه برو لس آنجلس ولی اگه دنبال تنوع بیشتر و یه دنیای عجیب و غریب از پیتزا با قیمت های مختلف میگردی برو نیویورک چون اونجا از پیتزای اقتصادی گرفته تا پیتزای لوکس پیدا میکنی ولی اگه فقط میوای یه پیتزا بخری و نمیخواهی ریاضی حل کنی سر قیمت لس آنجلس منطقی تره ولی اگه عاشق اینی وسط پونصد تا پیتزا فروشی گیر کنی و انتخاب سخت باشه برات نیویورک تو رو صدا میزنه !



رابطه ی بین امتیاز یک رستوران و قیمت آن:

```
ggplot(pizza_barstool, aes(x =  
as.factor(price_level), y = provider_rating)) +  
geom_violin(fill = "skyblue", color = "black") +  
labs(  
title = "توزیع امتیاز رستوران‌ها بر اساس سطح قیمت",  
x = "سطح قیمت",  
y = "امتیاز") +  
theme_minimal()
```

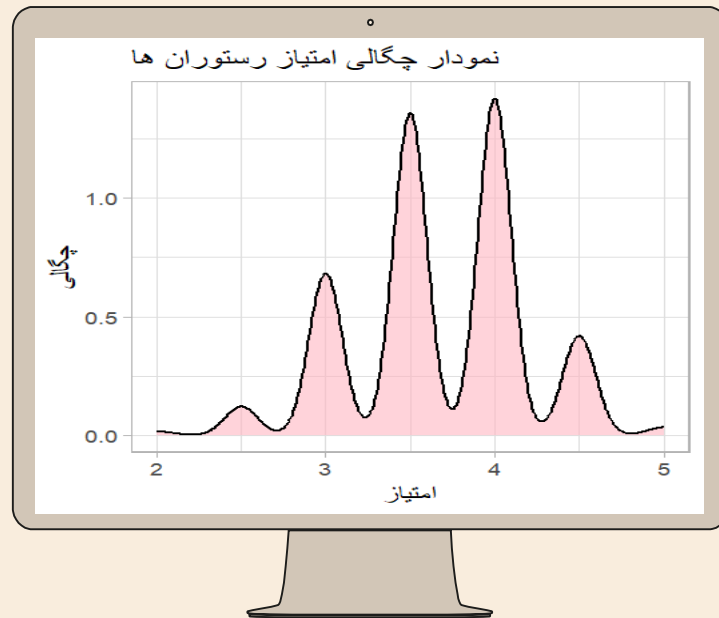


نتیجه گیری : طبق نمودار تو سطح قیمت صفر یا ارزون دامنه امتیاز ها زیاده و و چگالی بالایی توی اون دیده میشه که نشون میده بعضی از ارزون فروش ها واقعا دل میبرن و سطح قیمت یک و دو نوسانات زیاده یعنی ما پیتزای افتضاح تا عالی و توی این قسمت داریم و توی سطح سه تمرکز امتیاز ها بیشتره و میتونیم بگیم گرون فروشا اغلب استاندارد قابل قبولی دارند ولی ستاره نیستند !



بررسی امتیاز پیتزا ها :

```
ggplot(pizza_barstool, aes(x = provider_rating)) +  
  geom_density(fill = "lightpink", alpha = 0.6) +  
  labs(title = "نمودار چگالی امتیاز رستوران ها",  
        x = "امتیاز", y = "چگالی") +  
  theme_light()
```

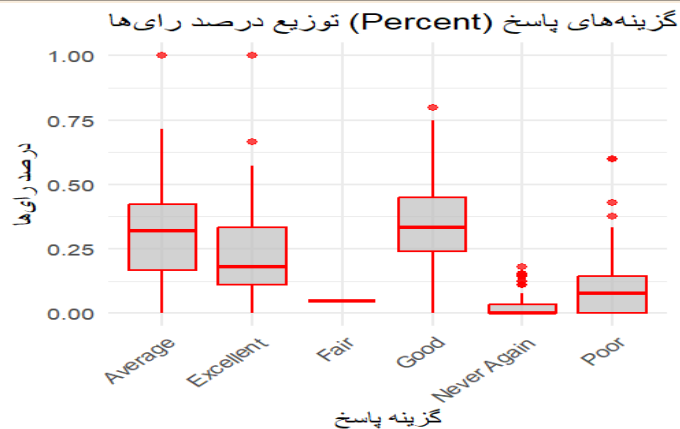


نتیجه گیری : طبق نمودار بیشتر تراکم امتیاز ها بین بازه ی سه و نیم تا چهار و نیم هست که این نشون میده اکثر پیتزا فروشی ها خوب تا خیلی خوب عمل کردن و چون ما سه قله ی مشخص داریم میتونیم بگیم سه گروه اصلی از نظر کیفیت داریم اما گروه کوچیکی هم امتیاز پایینی گرفتن که احتمالا اون گروه پیتزای فریز شده رو گرم میکنن و تحویل مردم میدن !

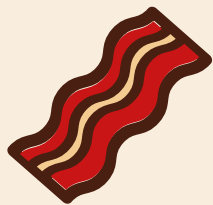


نمودار توزیع درصد رای های نظر سنجی راجع پیتزا :

```
ggplot(pizza_jared, aes(x = answer, y = percent)) +  
  geom_boxplot(fill = "gray", color = "red", alpha = 0.7) +  
  labs( title = "توزیع درصد رای ها (Percent) بر اساس گزینه های پاسخ (Answer)",  
        x = "گزینه پاسخ",  
        y = "درصد رای ها" ) +  
  theme_minimal() +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

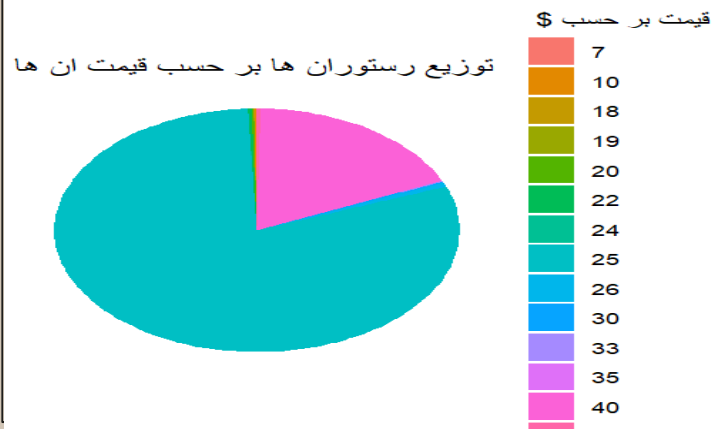


نتیجه گیری : بیشترین رای رو توی گزینه ها average آورده چون میانه ان بالاتره و excellent دامنه گسترده تری داره یعنی اختلاف نظر اینجا خیلی زیاده و گزینه های good/fair رای کمتری آوردن و اینجا نشون دهنده اینه که اکثر افراد یا خیلی راضی بودن یا خیلی ناراضی و تقریبا کمترین رای رو never again داره چون میانش نزدیک به صفره و تقریبا میتونیم این نظر سنجی شبیه به پیتزا پارتی بوده که اکثر مهمونا گفتن بد نبوده قابل تحمله و به عده هم گفتن عالیه و اونایی که تجربه بدی داشتن با تمام وجود گفتن never again / poor و با احساسات قوی رای دادن انگار پیتزا ها حسابی دل بردن یا دل شکستن !



نمودار توزیع رستوران های پیتزا فروشی بر اساس قیمت ان ها :

```
pizza_data %>%  
  group_by(price_range_max) %>%  
  summarise(count = n()) %>%  
  mutate(percent = count / sum(count) * 100) %>%  
  ggplot(aes(x = "", y = percent, fill = factor(Price_range_max))) +  
  geom_col(width = 1) +  
  coord_polar("y") +  
  labs(title = "توزیع رستوران ها بر اساس قیمت آن", fill = "$ قیمت بر حسب")  
  +  
  theme_void()
```

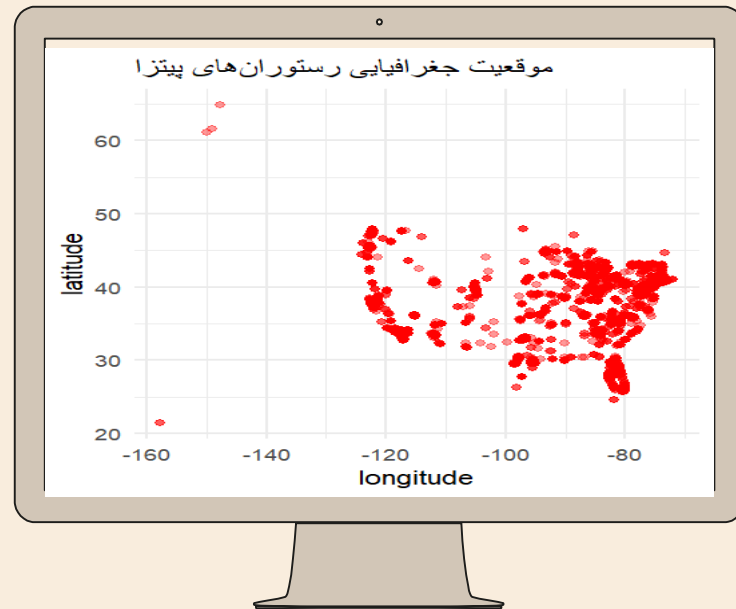


نتیجه گیری : طبق نمودار حدود هشتاد درصد پیتزا فروشی ها اکثرا توی بازه 24 تا 26 دلار فعالیت میکنن یعنی بیشتر مردم حاضرین برای یه پیتزای درست درمون یه مبلغ متوسط رو پرداخت کنن و بخش کوچیکی از بازار هم رفته سمت پیتزاهای گرون یعنی بالای 35 دلار / و خب پیتزاهای ارزون هم اونقدر کم هستنند که ادم شک میکنه که شاید بیشتر شیشه نون خشک با پنیر باشه تا یه پیتزای واقعی به خاطر همین کسی سمتش نمیره!



پراکندگی و موقعیت جغرافیایی رستوران های پیتزا :

```
ggplot(pizza_datafiniti, aes(x = Longitude, y = Latitude)) +  
  geom_point(alpha = 0.4, color = "red") +  
  labs(title = "موقعیت جغرافیایی رستوران های پیتزا") +  
  theme_minimal()
```



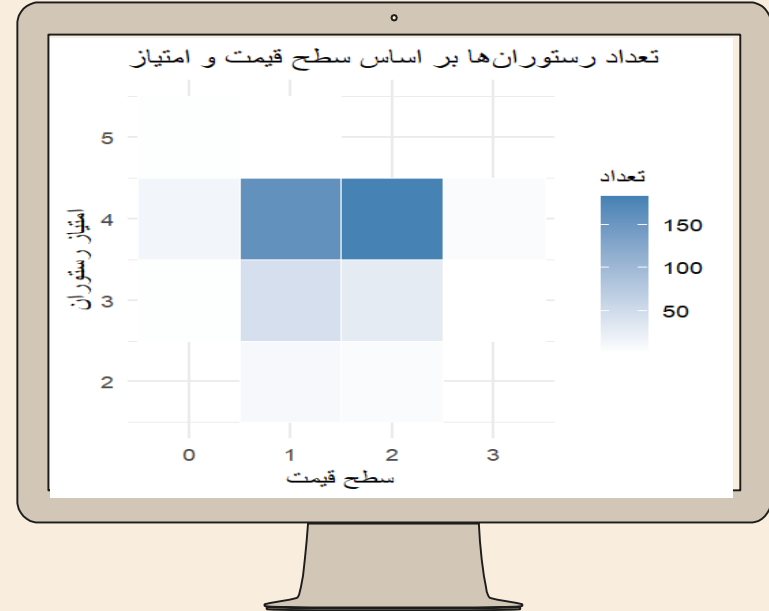
نتیجه گیری : نوار شرقی امریکا مثل نیویورک نیوجرسی و بوستون خیلی شلوغ و پرتراکم تراکم جمعیت بالاست و رقابت سنگینه و همه دنبال یه پیتزای درست درمون میگردن و مناطق بدون نقاط در واقع مناطق کوهستانی کم جمعیت و دور از مراکز تجاری هستن و دلایل نبود علاقه به پیتزا نیست پس اگه دنبال رقابت و پیشرفتی برو وسط نیویورک و سعی کن بین اون آسمان خراش ها بدرخشی یا برو نقطه ای که هنوز کسی نرفته و اولین ستاره ی پیتزا فروشی اونجا باش !





تعداد رستوران ها بر اساس قیمت و امتیاز ان ها :

```
pizza_barstool %>%  
  filter(!is.na(price_level), !is.na(provider_rating)) %>%  
  mutate(price_level = factor(price_level)) %>%  
  count(price_level, provider_rating = round(provider_rating))  
%>%  
  ggplot(aes(x = price_level, y = provider_rating, fill = n)) +  
  geom_tile(color = "white") +  
  scale_fill_gradient(low = "white", high = "steelblue") +  
  labs(title = "تعداد رستوران ها بر اساس سطح قیمت و امتیاز",  
       x = "سطح قیمت",  
       y = "امتیاز رستوران",  
       fill = "تعداد") +  
  theme_minimal()
```



نتیجه گیری : بیشتر رستوران ها دقیقاً جمع شدن وسط زمین بازی یعنی جایی که قیمت ها نه گرونه نه ارزون و امتیاز ها هم بین سه و چهار نوسان میکنه در واقع نه گرون فروشی جواب میده و نه غذای بی کیفیت / تعادل قیمت و کیفیت همون جاییه که مشتری ها براش صف کشیدن !



از توجه شما سپاسگزاریم!

داده های پیتزا به ما نشون دادند که قیمت بالا همیشه تضمین کننده ی کیفیت یا رضایت مشتری نیست و تنوع سلیقه ها سبک رستوران ها و حتی مکان های جغرافیایی میتواند نقش کلیدی در تجربه ی پیتزایی افراد داشته باشند و ما با کمک نمودار های اماری توانستیم الگو های جذاب و غیر منتظره را کشف کنیم و به زبان ساده تر توی دنیای پیتزا داده ها هم مثل پنیر کش میان و همیشه یه چیزی برای کشف کردن دارن

معصومه رنجبر / درس اصول و مصور سازی داده (استاد امین روشنی)