

Parte I-Pandas

Se deberán realizar las consignas pedidas sobre los datasets encontrados en la carpeta Dataset.

Ejercicio P4: Muestre el nombre del usuario que ha hecho una review al mismo lugar mas veces, y también el nombre del lugar.

Solución:

https://colab.research.google.com/drive/1IKYB7PlrtAnjMkFjQYXLRoUbzhVeJvN#scrollTo=l8f6m_uqHlpY

Ejercicio P10: ¿Cuál es el nombre del usuario que hizo más reviews en el estado de California?

Solución:

<https://colab.research.google.com/drive/1gP9bbYIJdQXFWgq7XpxkoiWn2y88K58m#scrollTo=0MBBOo17TQ8L>

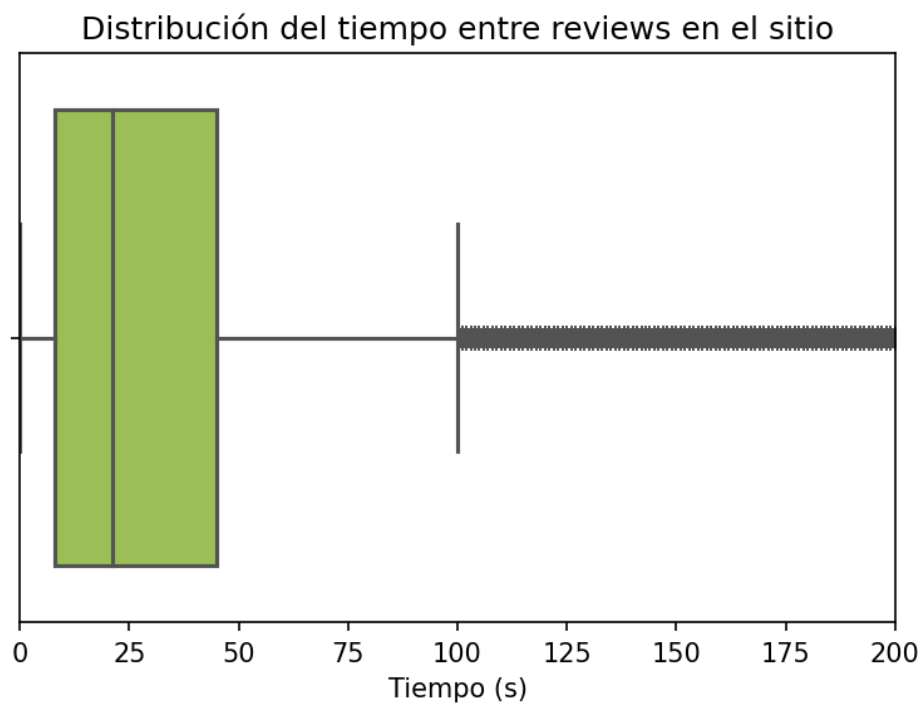
Ejercicio P23: ¿Cuál es la correlación entre el score que dejaron los usuarios en cada review después de 2017, y el score promedio que dejaron en todas las reviews del 2017? Calcule sin usar samples.

Solución: <https://colab.research.google.com/drive/1W0F-augyq9-iuCnDOR1irPI88stF3M4w>

Ejercicio N6: Uno de los mayores problemas de las reviews es que los usuarios tienen distintas escalas en puntaje para las mismas experiencias, entonces las estrellas que suele dejar un usuario no son comparables con las otro. Para normalizar un poco más esto, Yelp quiere que el usuario primero escriba el texto y según qué tan positivo o negativo sea sugerirle la cantidad de estrellas a dejar. vaderSentiment y textblob de python son dos librerías para analizar sentimientos, ambas pueden darnos un score de que tan positivo o negativo es el sentimiento de un texto. Para los textos reviews: ¿Cuál de los dos scores de emoción mejor correlaciona con el puntaje que esa review deja?

Solución: https://colab.research.google.com/drive/1ld3nbaxnXRaVhlw_Mbo4D3M_P8lyCgJS

Ejercicio V8: Consiga exactamente esta visualización utilizando los datos:

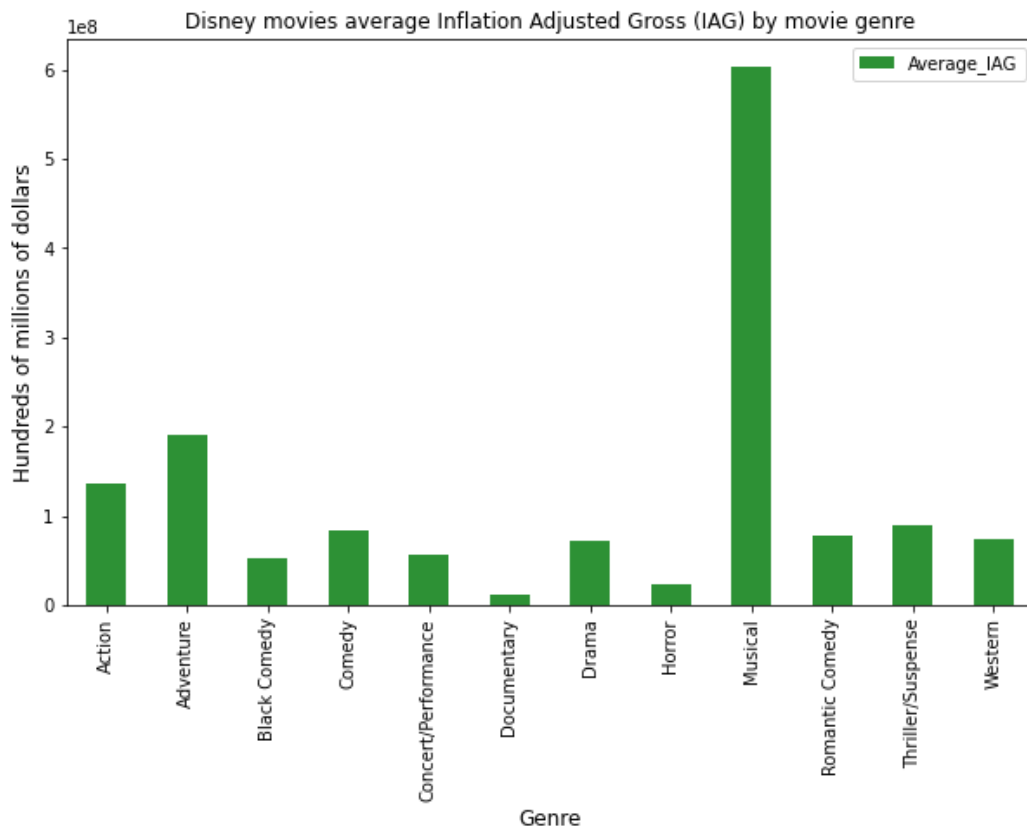


Solución: <https://colab.research.google.com/drive/1oiFo4enaoR-YBcUottFV-qRuSn-Qu0me#scrollTo=xAtcvEwV2lik>

Parte 2-Visualización de datos.

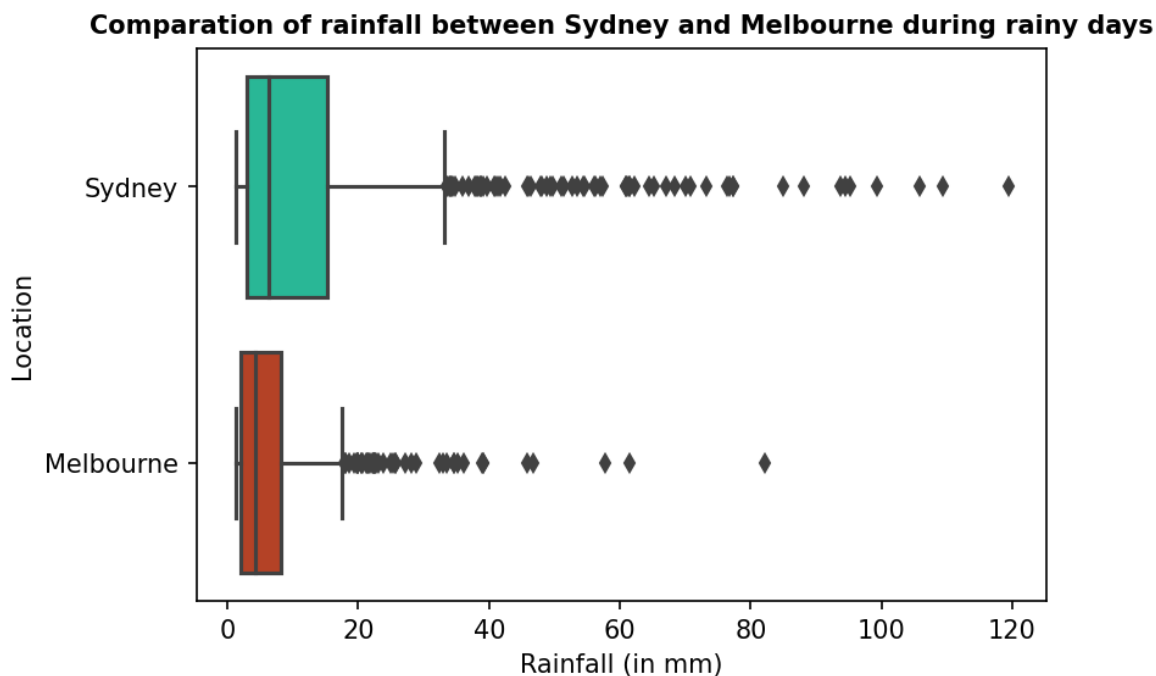
Realizar 6 visualizaciones con los datasets dados (se encontrarán en la carpeta Dataset). Se debe hacer al menos una visualización de los tipos: bar plot, density plot, violin plot, box plot, heatmap.

Visualización 1 (Bar plot): https://colab.research.google.com/drive/1ncSLfB-XchrJgJinT_LdPGkWhAXa-Pc#scrollTo=c_DQ3NyknTIg

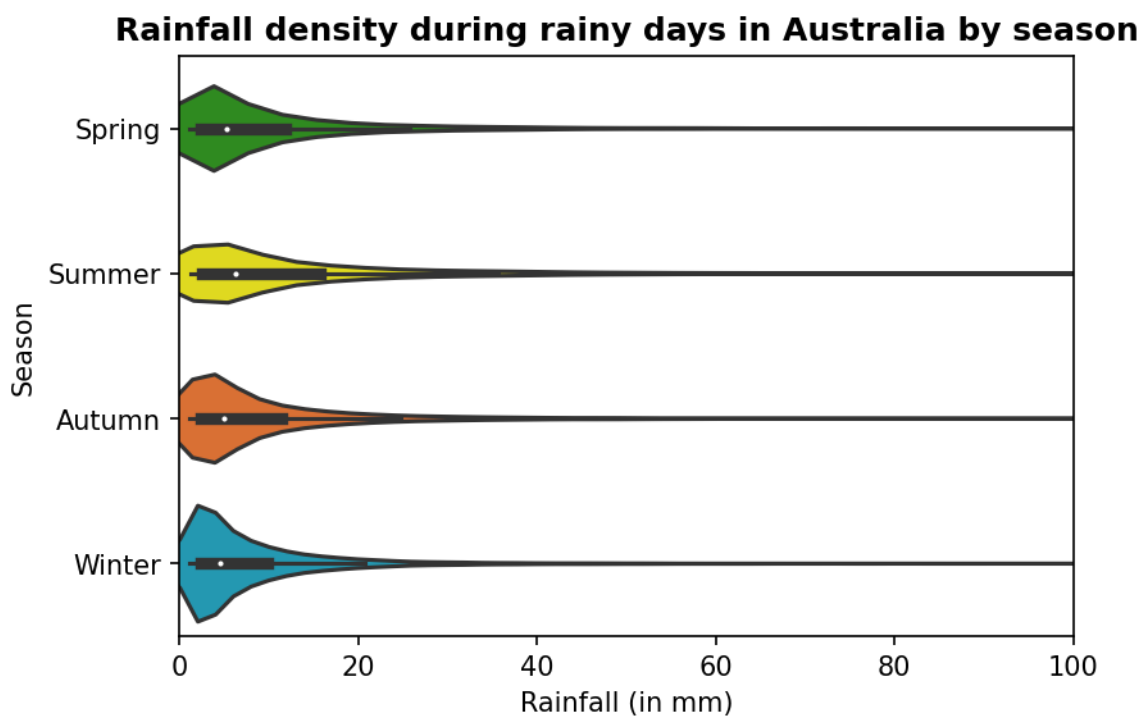


Visualización 2 (box plot):

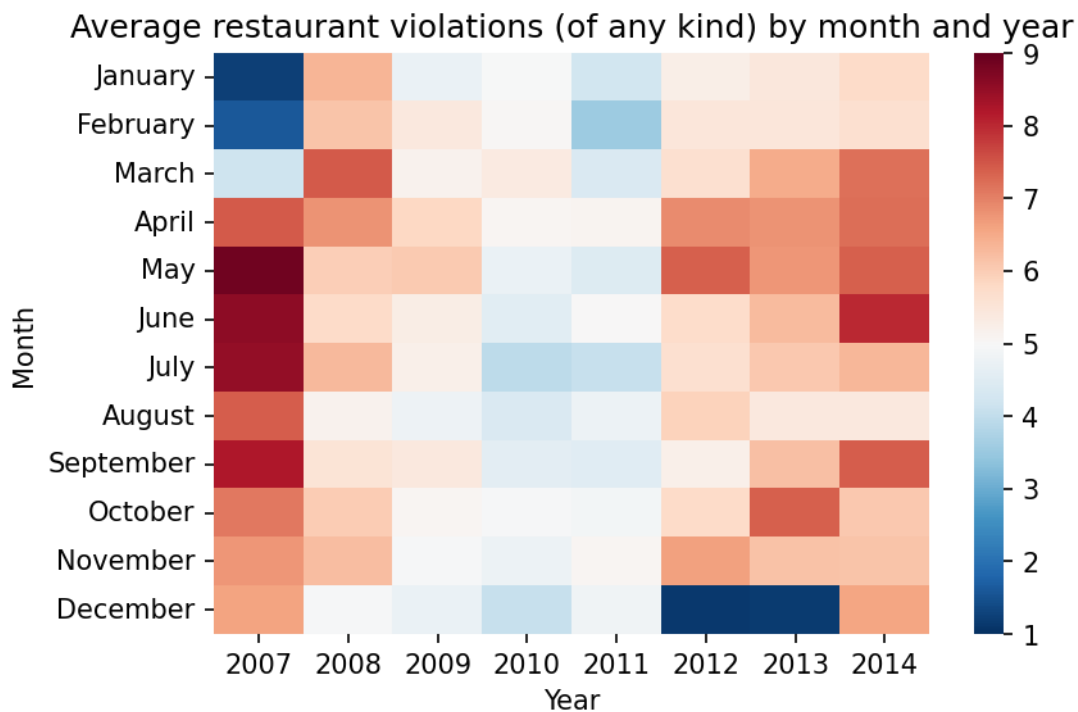
https://colab.research.google.com/drive/1OFrtzigV9maXb6GDDY5Nt8o_n2CAHt_b



Visualización 3 (violinplot): <https://colab.research.google.com/drive/1BcpVdeiFQRZwNjMQ-YszTFZ7GzmY144o#scrollTo=x9q2RSJERavO>

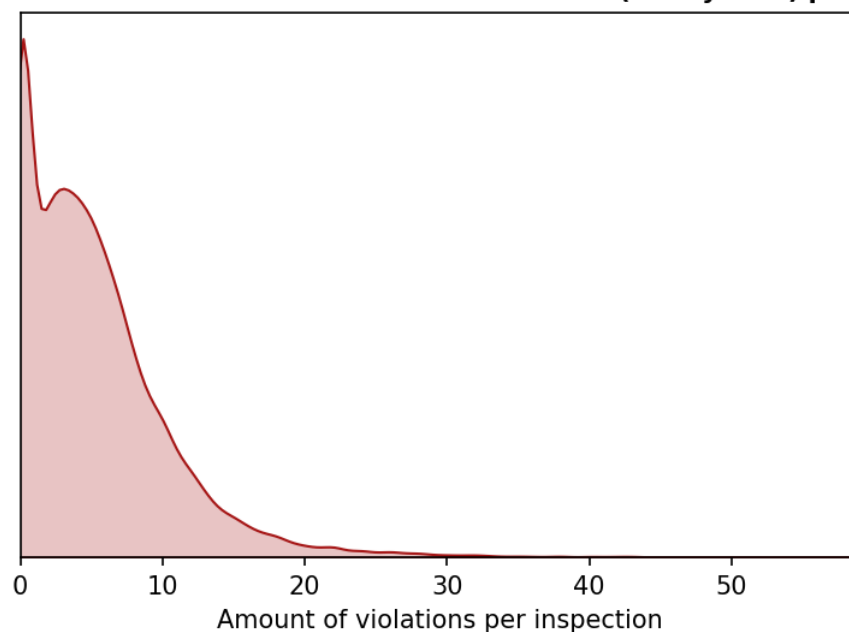


Visualización 4 (heat map): https://colab.research.google.com/drive/1Nzfx-H_AP-JBz2uJhIvSPdGKn8k8g6h



Visualización 5 (density plot): <https://colab.research.google.com/drive/1gYZdO3Ag4-Tw9j0TBmfBmiUw2LP20HY9#scrollTo=Fqq1QxsU4POu>

Distribution of total amount of restaurant violations (of any kind) per inspection



Visualización 6 (pie chart): <https://colab.research.google.com/drive/1dYqq0XoplR5FOCbOy5V6tZrvcejf8gmb>

Percentage of disney movies with over 100 million inflation adjusted gross (IAG)

