|  |  |
| --- | --- |
|  | **Task** |
| 1 | What is the overall average tip? |
| 2 | Get a numerical summary for 'tip' - are the median and mean very different? What does this tell you about the field? |
| 3 | Prepare a boxplot for 'tip', are there any outliers? |
| 4 | Prepare a boxplot for 'total\_bill', are there any outliers? |
| 5 | Gender: what is the percent of females in the data? |
| 6 | Prepare a bar plot with the bars representing the percentage of records for each gender. |
| 7 | Does the average tip differ by gender? Does one gender tip more than the other? |
| 8 | Does the average tip differ by the time of day? |
| 9 | Does the average tip differ by size (number of people at the table)? |
| 10 | Do smokers tip more than non-smokers? |
| 11 | Gender vs. smoker/non-smoker and tip size - create a 2 by 2 and get the average tip size. Which group tips the most? |
| 12 | Create a new metric called 'pct\_tip' = tip/ total\_bill - this would be percent tip give, and should be a better measure of the tipping behaviour. |
| 13 | Does pct\_tip differ by gender? Does one gender tip more than the other? |
| 14 | Does pct\_tip differ by size (number of people at the table)? |
| 15 | Make the gender vs. smoker view using pct\_tip - does your inference change? |
| 16 | Make a scatter plot of total\_bill vs. tip. |
| 17 | Make a scatter plot of total\_bill vs. pct\_tip. |