



# COMP 150: Pandas and Plotly

## Week 11, Lab 06



Setup:

1. Make a new colab notebook as well as a report (such as a Google document).
2. `import plotly.express as px`
3. Load the tips dataset: `df = px.data.tips()`

Answer the following, then upload both your report document as well as your colab notebook to Sakai:

4. Look at the data. In general, what is in this dataset?
5. What does the “time” column describe in the dataset?
6. What are all of the days represented in the dataset? (hint: use `unique()` on the “days” column)
7. Create a scatter plot. Let `x` be the `total_bill` and `y` be the `tip`. What is the highest tip/bill point? What is the lowest? Copy and paste this figure into your submission.
8. Create a box plot. Let `x` be the `day` and `y` be the `tip`. Which day has the highest average tipping, and what is the value? Copy and paste this figure into your submission.
9. Create another box plot. This time, add a `facet_row` to put tips from different “time” values on different rows. Which day has the highest vs the lowest average tipping? Copy and paste this figure into your submission.
10. Now, create another box plot. Add a “color” field to differentiate between sexes, and use “**boxmode=group**” to put like bars near each other. Use this figure to answer the question: do we have data for both sexes for all days of the week? Explain. You can use your figure.