# Lab- session 3

February 2, 2018

## 1 Lab: Session 3

### 1.0.1 Pandas Library

1. Import `pandas`, `numpy`, and `matplotlib` Python libraries.
2. Import the `sales_data.csv` dataset into Python. Make sure that the date is parsed correctly.
3. Create some summary statistics for the sales data.
4. Create a histogram to understand the distribution of purchasing patterns
5. Create a new dataset from the sales data, by having only `name`, `ext.price` columns for all transactions.
6. How many transactions did each customer have?
7. How much did each customer pay in total? Arrange the dataframe in increasing order.
8. Create a bar chart to show the distribution of total sales for all customers.
9. Create a table that shows the total sales for each customer by category.
10. Visualize the data in the previous question, using a stacked bar graph.
11. Create a line plot to check if we have certain months that are busier than others.

### 1.0.2 Pivot Tables

12. Import the `sales-funnel.xlsx` dataset into Python.
13. Create a pivot table that shows you the number of units sold by each rep for each product.
14. Create a pivot table that shows you the total quantity for each rep, and the max number of units per transaction
15. Create a pivot table that shows you the number of units sold by each rep for each product
16. Create a pivot table that shows you the number of units sold by each rep for each product with the margins of the table showing