MATH411 | Fall 2018 | Homework 2 (Due: Friday in class, 9/28/2018)

Your Name Here

Date Submitted Here

Problem 1 (Writing Function and Plotting Curve)

Imagine a monopolist selling a specific product with demand curve $Q(p)$, where $Q(p)$ is the quantity sold given a specific price p. To simplify things, let's suppose that $Q(p)$ is a linear function:

$$Q(p) = \alpha p + \beta$$

The total revenue will be given by:

$$R(p) = pQ(p) = \alpha p^2 + p\beta$$

(a) Code $R(p)$ in R by using $\alpha = -40$ and $\beta = 500$.

(b) Plot $R(p)$ VS p for p between 1 and 12. (Make your graph as nice as possible)

Problem 2 (Categorical Variable vs. Numerical Variable)

Amazon's new headquarters Scrape the table (i.e., the `Twenty-six cities` data) from the `cbsnews` website at https://www.cbsnews.com/news/amazon-hq2-cities-location-choices-new-second-h Tidy the data, then

(a) Print the First 5 and bottom 5 rows of your data.

(b) Make a bar plot to show the distribution of states in the data, rank the states by the number of cities in it from highest to lowest.

(c) Make a horizontal bar plot of `Percent with bachelor's degree` VS `Metro area`. Rank the `Metro area` by their `Percent with bachelor's degree` and label the percentage, i.e., %, on top of each `Metro area`.

Problem 3 (Categorical Variables and Numerical Variables)

Pittsburgh Penguins Scrape the Pittsburgh Penguins' `Team Record By Season` data from https://www.nhl.com/penguins/team/season-by-season-record.

(a) Seperate the last column `Finish` into two columns `rank` and `region`. (Hint: YOu can use the `separate` function from `tidyverse`). Delete the `NA` values in your data (Hint: you need to delete the 2004-05 season). Then print out the first 5 and bottom 5 rows.

(b) Make a bar plot to show the distribution of `rank` over all the seasons. Rank the `rank` by its frequency from highest to lowest.

(c) Plot the distribution of `GF` (Goals for, ie.e, goals scored by the Penguins).

(d) Make a new variable called `win.probability`, which can be calculated by $\frac{W}{GP}$ (i.e., numberf of game wins divided by number of game played). Plot the density distribution of `win.probability` and hightlight the mean of `win.probability` on the density plot as a big point.

(e) Make another variable called `GFpergame` (GF/GP). Then make a scatter plot between `win.probability` and `GFpergame`. Comment on the pattern you can tell from the scatter plot.

(f) Make another variable called `GApergame` (GA/GP). Then make a scatter plot between `win.probability` and `GApergame`. Comment on the pattern you can tell from the scatter plot.

(g) Make a scatter plot between `GFpergame` and `GApergame`. Comment on the pattern you can tell from the scatter plot.