**Exercises for Lecture 1**

Submission deadline: **March-26-2024**

1. Generate a sequence $(0, 1, \ldots, 30)$ with function `range` and assign it to a list `x`.

   a. Extract $(0, 1, \ldots, 10)$ from `x` with the slicing operator.

   b. Extract $(21, 22, \ldots, 30)$ from `x` with the slicing operator.

   c. Extract $(10, 12, \ldots, 24)$ from `x` with the slicing operator.

   d. Extract $(21, 23, \ldots, 29)$ from `x` with the slicing operator.

   e. Use list comprehension to assign $(31, 32, \ldots, 40)$ in a list `y`.

   f. With list `y`, make `x` be a list of numbers $(0, 1, \ldots, 40)$

   g. Remove the number 33 from `x` with method `remove`.

   h. Delete numbers $(31, 32, \ldots, 40)$ in `x` with function `del`.

2. Create a dictionary looks like:

| Team | Points | Result |
|---|---|---|
| Nuggets | 94 | Win |
| Heat | 89 | Lose |

   a. Extract Team and Points from the dictionary and print them.

   b. Extract the first element in Team from the dictionary and print it.

   c. Extract the second element in Points from the dictionary and print it.

   d. Change Points to $(108, 111)$ and Result to (Lose, Win) and print the changed dictionary.

3. Import two data files: `SP500m.csv` and `NASDAQm.csv` with `pandas` function `read_csv`.

   a. Transfer dates in column `Date` to ISO-8601 format with method `to_datetime`.

   b. With adjusted close prices (data in column `Adj Close`), calculate monthly log returns: $\log(P_t) - \log(P_{t-1})$, where $P_t$ and $P_{t-1}$ are adjusted close prices at period $t$ and $t-1$. Assign the log returns into a new column called `logr`. Note that the first element in this new column should be `nan`.

   c. With adjusted close prices, calculate monthly simple returns: $(P_t - P_{t-1})/P_{t-1}$. Assign the simple returns into a new column called `sr`. Note that the first element in this new column should be `nan`.

d. Use method `describe` to calculate summary statistics of the data and assign the results of the summary statistics into `result1` and `result2`.

e. Use functions `skew` and `kurtosis` in `scipy` to calculate skewness and kurtosis of the data. Before you calculate, you need to drop the column `Date`. Then you need to use method `apply`. Also `nan` should be omitted in the calculations.

f. Combine results of skewness and kurtosis with those in `result1` and `result2`. To do this, at first you may need to transfer results of skewness and kurtosis to `pandas` data frame with a column name. Then use function `concat` to combine these results.

g. Output these results with method `to_csv`. The names of the output files are `result1.csv` and `result2.csv`. Note that the output should keep the row index. Also use the representation of `NaN` for not available number by setting `na_rep = 'NaN'` in the method.

h. Print out sample correlation between log and sample returns.

4. a. Plot probability mass function of a binominal distribution with parameter $n = 10$ and $p = 0.2$. Save the plot with method `savefig` and file name `figure1.png`.

b. Plot probability density function of a normal distribution with mean equal to 0.5 and standard deviation equal to 1.2. Save the plot with method `savefig` and file name `figure2.png`.

5. Conduct a simulation to verify unbiasedness.

a. Use `random.seed` to initialize the random number generator.

b. What are mean and variance of a random variable following exponential distribution with parameter $\lambda = 1.2$.

c. Generate 100 random samples from an exponential distribution with parameter $\lambda = 1.2$. Calculate sample mean and sample variance.

d. Repeat question c. again and compare the calculated sample mean and sample variance with those obtained in question c.

e. Repeat question c. 10,000 times and collect the sample mean and sample variance estimates.

f. Calculate sample means of the 10,000 sample mean and sample variance estimates and compare them to values of the population mean and variance.