



# Drinking Habits of Global Population

## Objective

At the end of this project you will be able to:

- Showcase your data analysis learning, applying all the concepts you have learnt throughout the course.

## Instructions

New York is a global city with over 846 spoken languages. A new wine and brewery establishment want to analyze the drinking habits of the population by nationality to help determine their inventory. As a Data Scientist, you are asked to lead this project, and analyze the data of alcohol consumption around the world to better serve your company's diverse customer base.

In this assignment, you will analyze and predict alcohol consumption by country using attributes or features such as beer servings and wine servings. A template notebook is provided in the lab. Some hints to the questions are given in the template notebook. You will use Watson Studio to perform the analysis and share an image of your finished notebook with a URL.

## Review Criteria

**There are ten questions to this final assignment. This project is worth 20% of your total grade. This final project will be graded by your peers who are also completing the course during the same session.**

1. Display the data types of each column using the attribute dtype.
2. Use the method groupby to get the number of wine servings per continent.
3. Perform a statistical summary and analysis of beer servings for each continent.
4. Use the function boxplot in the seaborn library to produce a plot that can be used to show the number of beer servings on each continent.
5. Use the function regplot in the seaborn library to determine if the number of wine servings negatively or positively correlated with the number of beer servings.
6. Fit a linear regression model to predict the 'total\_litres\_of\_pure\_alcohol' using the number of 'wine\_servings' then calculate R2.
7. Use the list of features to predict the 'total\_litres\_of\_pure\_alcohol', split the data into training and testing and determine the R2 on the test data, using the provided code.
8. Create a pipeline object that scales the data, performs a polynomial transform and fits a linear regression model. Fit the object using the training data in the question above, then calculate the R 2 using the test data. Take a screenshot of your code and the R2. There are some hints in the notebook.
9. Create and fit a Ridge regression object using the training data, setting the regularization parameter to 0.1 and calculate the R2 using the test data. Take a screenshot of your code and the R2.
10. Perform a 2nd order polynomial transform on both the training data and testing data. Create and fit a Ridge regression object using the training data, setting the regularization parameter to 0.1. Calculate the R2 utilizing the test data provided. Take a screen-shot of your code and the R2.

**For each item, you should take a screenshot of the output of each cell with the code that generated it. The provided code must be run. Most of the questions are independent, so if you miss one you can still complete the remaining problems. You will also be asked to share your notebook. You will receive marks for displaying your notebook.**

**Step A:** Create an account in Watson Studio if you don't have an account already. (If you already have an account, jump to Step B).

1. [Final Project Setup](#)

**Step B:** Sign into Watson Studio and import your notebook

1. Sign in to [IBM Cloud Watson Studio](#)
2. Click on "Create a project".
3. On the Create a project page, click "Create an empty project".
4. Give a name to your project and a description for your reference, then set-up your project as follows, then click "Create".

**Notice 1:** Because you are going to share this project with your peers for evaluation, please make sure you uncheck "Restrict who can be a collaborator".

**Notice 2:** You must create an IBM Object Storage, if you don't have an IBM Object Storage, you can use the free Lite plan.

5. From the top-right, click on "Add to project" and then select "Notebook".
6. In the "New Notebook" form, click on "From URL" and right click the [Notebook URL](#), copy the link and paste it.
7. Give your notebook a proper name and description and click on "Create Notebook" to initialize the notebook.

**Step C:** Complete the Notebook

1. Start running the notebook.
2. Complete the notebook based on the description in the notebook.

**Step D:** Share the Notebook

1. Click on the share icon on the top-right side of your page.
2. Activate the "Share with anyone who has the link".
3. Select "All content excluding sensitive code cells".
4. Copy the link from "Permalink to view notebook".

## Submit your Notebook for Grading

Paste the shared link of your Notebook in the provided text box below for peer-review.