



[Course](#) > [Modul...](#) > [Final A...](#) > Peer-G...

Peer-Graded Assignment: Drinking Habits of Global Population

New York is a global city with over 846 spoken languages. A new wine and brewery establishment wants 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.

There are ten questions to this final assignment. Simple questions are worth one mark and more difficult questions are worth two marks. For each item, you should take a screen shot 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.

This assignment will be graded by your peers who are also completing this course during the same session.

OPEN RESPONSE ASSESSMENT

Status

This assignment is in progress. Complete the learner training step to move on.

▼ **Your Response** due Aug 30, 2020 17:00 PDT (in 1 month) ✓ COMPLETE

Status

Your response has been submitted. You will receive your grade after all steps are complete and your response is fully assessed. You still need to complete the peer assessment and self assessment steps.

The question for this section

Place the URL of your notebook in here as per the following instruction [How to share you lab](#)

Your response

https://dataplatform.cloud.ibm.com/analytics/notebooks/v2/7e52e885-3b7e-4925-85e7-b552810ea3af/view?access_token=ebc064f8bed416b47fd02017bc803cd432f8ab00d77bacf54eb9eda5ea1b86f4

Your Uploaded Files

Files that were uploaded by you:
1 ([Final Assignment .pdf](#))

Learn to Assess Responses IN PROGRESS (2 OF 2) due Aug 30, 2020 17:00 PDT (in 1 month)

Learning to Assess Responses

Your assessment differs from the instructor's assessment of this response. Review the response and consider why the instructor may have assessed it differently. Then, try the assessment again.

The question for this section

Place the URL of your notebook in here as per the following instruction [How to share you lab](#)

The response to the prompt above:

Replace this text with your own sample response for this assignment. Then, under Response Score to the right, select an option for each criterion. Learners practice performing peer assessments by assessing this response and comparing the options that they select in the rubric with the options that you specified.

▼ Did the user Display the data types of each column using the attribute dtype , as follows:

```
df.dtypes
country      object
beer_servings  int64
spirit_servings int64
wine_servings int64
total_litres_of_pure_alcohol float64
continent     object
dtype: object
```

Selected Options Agree

The option you selected is the option that the instructor selected.

☐ correct

was there the following output and code:

```
df.dtypes
```

```
country      object
beer_servings  int64
spirit_servings int64
```

```
wine_servings      int64
total_litres_of_pure_alcohol  float64
continent           object
dtype: object
```

1 POINTS

☒ **incorrect**

the output did not match

0 POINTS

- ▼ Did the user use the method `groupby` to get the number of wine servings per continent as follows:

```
df.groupby('continent')['wine_servings'].sum()
```

```
continent
Africa      862
Asia        399
Europe     6400
North America  564
Oceania      570
South America  749
```

```
Name: wine_servings, dtype: int64
```

Selected Options Agree

The option you selected is the option that the instructor selected.

☐ **correct**

Does the user have the following output:

```
continent
Africa      862
Asia        399
Europe     6400
North America  564
Oceania      570
South America  749
```

```
Name: wine_servings, dtype: int64
```

1 POINTS

☒ incorrect

the result did not match.

0 POINTS

- ▼ Did the user Perform a statistical summary and analysis of beer servings with the following line of code:
- ```
df.groupby('continent')['beer_servings'].describe()
```

View this url for the output (please paste in web browser): [https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final\\_proget\\_images/Question\\_3\\_table.png](https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final_proget_images/Question_3_table.png)

---

### Selected Options Agree

---

The option you selected is the option that the instructor selected.

☐ correct

yes, the user had the correct code and/or output:

```
df.groupby('continent')['beer_servings'].describe()
```

View this url for the output (please paste in web browser): [https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final\\_proget\\_images/Question\\_3\\_table.png](https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final_proget_images/Question_3_table.png)

1 POINTS

☒ incorrect

the output or code was incorrect

0 POINTS

- ▼ The user use the following lines of code to display a boxplot.

```
df.boxplot(column = 'beer_servings', by = 'continent')
plt.show()
```

or/and or the following image (please paste in web browser) : [https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final\\_project\\_images/Question\\_4\\_table.png](https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final_project_images/Question_4_table.png)

---

### Selected Options Agree

---

The option you selected is the option that the instructor selected.

☐ correct

the user displayed the correct code and/or image:

```
df.boxplot(column = 'beer_servings', by = 'continent')
plt.show()
```

or/and or the following image (please paste in web browser) : [https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final\\_project\\_images/Question\\_4\\_table.png](https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final_project_images/Question_4_table.png)

1 POINTS

☒ incorrect

the code and output was incorrect.

0 POINTS

- ▼ Did the user use the function regplot in the seaborn library to determine if the number of wine servings is negatively or positively correlated with the number of beer servings, with the following output:

```
sns.regplot(x = "beer_servings", y = "wine_servings", data = df)
plt.show()
```

or/and or the following image (please paste in web browser) : [https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final\\_project\\_images/Question\\_5\\_table.png](https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final_project_images/Question_5_table.png)

---

### Selected Options Agree

---

The option you selected is the option that the instructor selected.

☐ correct

the user displayed the following line of code and/ or the following image:

```
sns.regplot(x = "beer_servings", y = "wine_servings", data = df)
```

```
plt.show()
```

or/and or the following image (please paste in web browser) : [https://s3-api.us-](https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final_progect_images/Question_5_table.png)

[geo.objectstorage.softlayer.net/cf-courses-](https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final_progect_images/Question_5_table.png)

[data/CognitiveClass/DA0101EN/edx/project/Final\\_progect\\_images/Question\\_5\\_table.png](https://s3-api.us-gio.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/DA0101EN/edx/project/Final_progect_images/Question_5_table.png)

1 POINTS

☒ incorrect

the question is incorrect

0 POINTS

▼ The user Fit a linear regression model to predict the 'total\_litres\_of\_pure\_alcohol' using the number of 'wine\_servings' then calculated the

---

### Selected Options Agree

The option you selected is the option that the instructor selected.

☐ correct

The code and output should be:

```
lm = LinearRegression()
```

```
Y = df[['total_litres_of_pure_alcohol']]
```

```
X = df[['wine_servings']]
```

```
lm.fit(X, Y)
```

```
lm.score(X, Y)
```

```
Output:.4456875459787605
```

1 POINTS

☒ incorrect

the output was incorrect

0 POINTS

- ▼ Use the list of features to predict the 'total\_litres\_of\_pure\_alcohol', split the data into training and testing and determine the
- 

### Selected Options Agree

---

The option you selected is the option that the instructor selected.

☐ correct

the user had the following output

```
features =['beer_servings','spirit_servings','wine_servings']
X=df[features]
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.1, random_state=0)
lm = LinearRegression()
lm.fit(X_train, y_train)
lm.score(X_test, y_test)
output:0.6990304512837944
```

1 POINTS

☒ incorrect

had no output

0 POINTS

- ▼ Did the user Create a pipeline object that scales the data, performs a polynomial transform and fits a linear regression model? Then Fit the object using the training data in the question above, then calculate the  $R^2$  using the test data. The correct output and code is shown below.

```
Input=[('scale',StandardScaler()),('polynomial', PolynomialFeatures(include_bias=False)),
 ('model',LinearRegression())]
pipe=Pipeline(Input)

pipe.fit(X_train, y_train)
pipe.score(X_test, y_test)

output: 0.6990304512837944
```

---



### Selected Options Differ

---

The option you selected is not the option that the instructor selected.

☐ correct

there was the following output:

```
Input=[('scale',StandardScaler()),('polynomial', PolynomialFeatures(include_bias=False)),
('model',LinearRegression())]
pipe=Pipeline(Input)
```

```
pipe.fit(X_train, y_train)
pipe.score(X_test, y_test)
```

output: 0.6990304512837944

1 POINTS

☒ incorrect

0 POINTS

▼ Did the user create and fit a Ridge regression object using the training data, setting the regularization parameter to 0.1 and calculate the

---

### Selected Options Differ

---

The option you selected is not the option that the instructor selected.

☒ correct

The result is:

```
RR= Ridge(alpha=0.1)
RR.fit(X_train, y_train)
RR.score(X_test,y_test)
output:0.699
```

1 POINTS

☐ incorrect

the output is incorrect

0 POINTS

- ▼ Did the user 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
- 

### Selected Options Agree

---

The option you selected is the option that the instructor selected.

☐ correct

The output should be:

```
pr=PolynomialFeatures(degree=2)
x_train_pr=pr.fit_transform(X_train)
x_test_pr=pr.fit_transform(X_test)
RR= Ridge(alpha=0.1)
RR.fit(x_train_pr, y_train)
RR.score(x_test_pr,y_test)
```

output: 0.7076

1 POINTS

☒ incorrect

the output is incorrect

0 POINTS

Compare your selections with the instructor's selections

Assess Peers due Aug 30, 2020 17:00 PDT (in 1 month) NOT AVAILABLE

Assess Your Response due Aug 30, 2020 17:00 PDT (in 1 month) NOT AVAILABLE

► Your Grade: Not Completed

