Try again once you are ready
TO PASS 80% or higher

Try again

GRADE
43.33%

Simple Linear Regression: Graded Quiz

LATEST SUBMISSION GRADE
43.33%

1. Which of the following evaluation metrics can be used to evaluate a model while modeling a continuous output variable? **0.333 / 1 point**

- [x] Precision

This should not be selected
Incorrect. Please review Task 9 of this course. Hint: Evaluation metrics for classification problems, such as **Precision**, are not useful for regression problems. We need an evaluation metric that compares continuous values.

- [x] Accuracy

This should not be selected
Incorrect. Please review Task 9 of this course. Hint: Evaluation metrics for classification problems, such as **accuracy**, are not useful for regression problems. We need an evaluation metric that compares continuous values.

- [x] AUC-ROC

This should not be selected
Incorrect. Please review Task 9 of this course. Hint: Evaluation metrics for classification problems, such as **AUC-ROC**, are not useful for regression problems. We need an evaluation metric that compares continuous values.

- [x] Mean Squared Error

Correct
Correct! Linear Regression outputs continuous values. So we need an evaluation metric that compares continuous values. **Mean Squared Error** (MAE) is the mean of the absolute value of the errors:

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- [x] Logloss

This should not be selected
Incorrect. Please review Task 9 of this course. Hint: Evaluation metrics for classification problems, such as **logloss**, are not useful for regression problems. We need an evaluation metric that compares continuous values.

- [x] Mean Absolute Error

Correct
Correct! Linear Regression outputs continuous values. So we need an evaluation metric that compares continuous values. **Mean Absolute Error** (MAE) is the mean of the absolute value of the errors:

$$\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

$i=1$

2. ***true*** and ***pred*** are two lists of true and predicted response values from a simple linear regression model. Using scikit-learn, how would you calculate the root mean squared error to evaluate the response values? **0 / 1 point**

- ```
  np.sqrt(metrics.mean_squared_eror(true, pred))
  ```
- ```
  metrics.mean_squared_eror(true, pred)
  ```
- ```
  metrics.mean_absolute_error(true, pred)
  ```

Incorrect

Incorrect. This is the mean squared error (MSE) evaluation metric:

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

3. For a pandas DataFrame **df**, what function is used to return the first five rows of the DataFrame? **1 / 1 point**

- ```
  df.tail()
  ```
- ```
  df.head()
  ```
- ```
  df.info()
  ```

Correct

Correct! This function, by default, returns the first five rows of the DataFrame. It is useful for quickly testing if your object has the right type of data in it. It can also return the first *n* rows when specified as an argument. Eg: The code below returns the first 10 rows from **df**.

```
df.head(10)
```

4. A DataFrame, **df**, contains three columns: TV, Radio, Newspaper. Which of the following options drops the Radio column inplace? **2 / 2 points**

- ```
  df.drop(['TV','Newspaper'], axis=1, inplace=True)
  ```
- ```
  df.drop(['Radio'], axis=1, inplace=True)
  ```
- ```
  df.drop(['Radio'], axis=0, inplace=True)
  ```
- ```
  df.drop(['Radio'], axis=0, inplace=False)
  ```

Correct

Good job! The first argument specifies the columns to be dropped. The *axis=1* argument drops the columns

from **df**. Lastly, when *inplace=True*, the operation is performed inplace.

5. Which of the following Seaborn functions did we use to plot a univariate distribution of observations? **0 / 1 point**

```
sns.distplot()
```

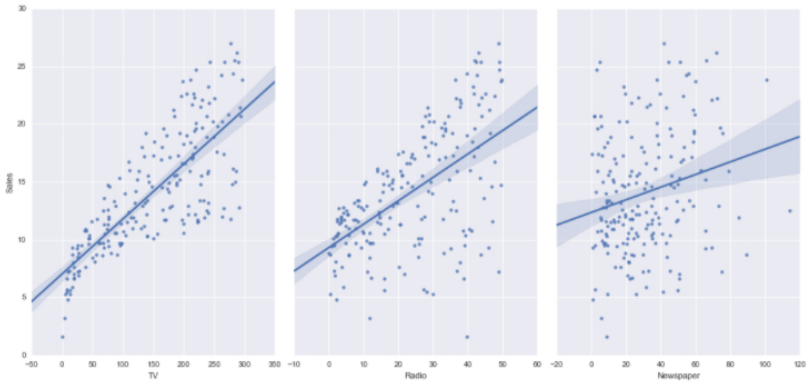
```
sns.kdeplot()
```

```
sns.jointplot()
```

Incorrect

Incorrect. Please review Task 4 of this course. This function is used to fit and plot a univariate or bivariate kernel density estimate.

6. The image below visualizes the relationship between the features *TV, Radio, Newspaper* and the response, *Sales*, using scatter plots. Each scatter plot is fit with a linear regression model. **1 / 1 point**



Select all the conclusions that can be drawn from the above plots.

- [x] There is a strong relationship between TV ads and Sales.

Correct

Correct! The plot on the left shows a strong, positive, linear association between TV ads and Sales with a few outliers.

- [x] There is a weak relationship between Radio ads and Sales.

Correct

Correct! There is a weak, positive, linear association between Radio ads and Sales with many outliers.

- [x] There is a very weak to no relationship between Newspaper ads and Sales

Correct

Correct! There is a weak, non linear relationship between Newspaper ads and Sales.

7. There are multiple features (explanatory variables) in a simple linear regression model. **0 / 1 point**

- ( ) False
- (•) True

Incorrect

Incorrect. Please review Task 6 of this course.

8. For the simple linear regression model below, what terms correspond to the feature coefficient and y-intercept **0 / 1 point**

respectively?

$$y = \beta_0 + \beta_1 \times TV$$

- ○ β_1 is the feature coefficient and β_0 is the y-intercept
- ◉ β_1 is the feature coefficient and y is the y-intercept
- ○ β_0 is the feature coefficient and β_1 is the y-intercept

Incorrect

Incorrect. Please review Task 6 and Task 7 of this course.

9. The first line of code below shows some data being split into training and test sets for features *X* and response *y*. What two objects returned from the *train_test_split* function need to be passed as arguments to *linreg.fit*( )?

0 / 1 point

```
1  X_train, X_test, y_train, y_test = train_test_split(X, y)
2
3  from sklearn.linear_model import LinearRegression
4
5  linreg = LinearRegression()
6  linreg.fit()
```

- ○ *X_train* and *X_test*
- ◉ *X_test* and *y_test*
- ○ *X_train* and *y_train*

Incorrect

Incorrect. Please review Task 6 and Task 8 of this course.