Module 9 Challenge

Start Assignment

Due Jan 30 by 11:59pm **Points** 100 **Submitting** a text entry box or a website url

Background

W. Avy likes your analysis, but he wants more information about temperature trends before opening the surf shop. Specifically, he wants temperature data for the months of June and December in Oahu, in order to determine if the surf and ice cream shop business is sustainable year-round.

What You're Creating

This new assignment consists of two technical analysis deliverables and a written report. You will submit the following:

- Deliverable 1: Determine the Summary Statistics for June
- Deliverable 2: Determine the Summary Statistics for December
- Deliverable 3: A written report for the statistical analysis (README.md)

Files

Use the following link to download the Challenge starter code.

Download challenge starter code (https://2u-data-curriculum-

team.s3.amazonaws.com/dataviz-online/module_9/SurfsUp_Challenge_starter_code.ipynb)

Deliverable 1: Determine the Summary Statistics for June (40 points)

Deliverable 1 Instructions

Using Python, Pandas functions and methods, and SQLAlchemy, you'll filter the date column of the Measurements table in the `hawaii.sqlite` database to retrieve all the temperatures for the month of June. You'll then convert those temperatures to a list, create a DataFrame from the list, and generate the summary statistics.

REWIND

For this deliverable, you've already done the following in this module:

- **Lesson 9.2.1:** Filter a `sqlite` table on a value from a column
- **Lesson 9.2.1:** Save query results as a DataFrame
- **Lesson 9.2.5:** Generate summary statistics

Follow the instructions below to complete Deliverable 1.

1. Download the `SurfsUp_Challenge_starter_code.ipynb` file into your surfs_up folder, then rename it `SurfsUp_Challenge.ipynb`.
2. Use the instructions below to add code where indicated by the numbered comments in the starter code file. The starter code file includes all dependencies needed for this Challenge.
3. In Step 1, write a query that filters the `date` column from the `Measurement` table to retrieve all the temperatures for the month of June.
4. In Step 2, convert the June temperatures to a list.
5. In Step 3, create a DataFrame from the list of temperatures for the month of June.
6. In Step 4, generate the summary statistics for the June temperatures DataFrame.
7. After you run Step 4 in your `SurfsUp_Challenge.ipynb` file, confirm that the summary statistics match the image below.

	June Temps
count	1700.000000
mean	74.944118
std	3.257417
min	64.000000
25%	73.000000
50%	75.000000

50%	75.000000
75%	77.000000
max	85.000000

Deliverable 1 Requirements

You will earn a perfect score for Deliverable 1 by completing all requirements below:

- A working query is written to retrieve the June temperatures from the `date` column of the `Measurement` table. **(10 pt)**
- The temperatures are added to a list. **(10 pt)**
- The list of temperatures is converted to a Pandas DataFrame. **(10 pt)**
- Summary statistics are generated for the DataFrame. **(10 pt)**

Deliverable 2: Determine the Summary Statistics for December (40 points)

Deliverable 2 Instructions

Using Python, Pandas functions and methods, and SQLAlchemy, you'll filter the date column of the Measurements table in the `hawaii.sqlite` database to retrieve all the temperatures for the month of December. You'll then convert those temperatures to a list, create a DataFrame from the list, and generate the summary statistics.

REWIND

For this deliverable, you've already done the following in this module:

- **Lesson 9.2.1:** Filter a `sqlite` table on a value from a column
- **Lesson 9.2.1:** Save query results as a DataFrame
- **Lesson 9.2.5:** Generate summary statistics

Follow the instructions below to complete Deliverable 2.

1. Use the instructions below to add code where indicated by the numbered comments in your `SurfsUp_Challenge.ipynb` file.
2. In Step 6, write a query that filters the `date` column from the `Measurement` table to retrieve all the temperatures for the month of December.
3. In Step 7, convert the December temperatures to a list.
4. In Step 8, create a DataFrame from the list of temperatures for the month of December.
5. In Step 9, generate the summary statistics for the December temperatures DataFrame.
6. After you run Step 9 in your `SurfsUp_Challenge.ipynb` file, confirm that the summary statistics match the image below.

	December Temps
count	1517.000000
mean	71.041529
std	3.745920
min	56.000000
25%	69.000000
50%	71.000000
75%	74.000000
max	83.000000

Deliverable 2 Requirements

You will earn a perfect score for Deliverable 2 by completing all requirements below:

- A working query is written to retrieve the December temperatures from the `date` column of the `Measurement` table **(10 pt)**
- The temperatures are added to a list. **(10 pt)**
- The list of temperatures is converted to a Pandas DataFrame. **(10 pt)**
- Summary statistics are generated for the DataFrame. **(10 pt)**

Deliverable 3: A written report for the statistical analysis (20 points)

Deliverable 3 Instructions

For this part of the Challenge, write a report that describes the key differences in weather between June and December and two recommendations for further analysis.

The analysis should contain the following:

1. **Overview of the analysis:** Explain the purpose of this analysis.
2. **Results:** Provide a bulleted list with three major points from the two analysis deliverables. Use images as support where needed.
3. **Summary:** Provide a high-level summary of the results and two additional queries that you would perform to gather more weather data for June and December.

Deliverable 3 Requirements

Structure, Organization, and Formatting (6 points)

The written analysis has the following structure, organization, and formatting:

- There is a title, and there are multiple sections. **(2 pt)**
- Each section has a heading and subheading. **(2 pt)**
- Links to images are working and displayed correctly. **(2 pt)**

Analysis (14 points)

The written analysis has the following:

1. Overview of the statistical analysis:
    - The purpose of the analysis is well defined. **(3 pt)**
2. Results:
    - There is a bulleted list that addresses the three key differences in weather between June and December. **(6 pt)**
3. Summary:
    - There is a high-level summary of the results and there are two additional queries to perform to gather more weather data for June and December. **(5 pt)**

Submission

Once you're ready to submit, make sure to check your work against the rubric to ensure you are meeting the requirements for this Challenge one final time. It's easy to overlook items when you're in the zone!

As a reminder, the deliverables for this Challenge are as follows:

- Deliverable 1: Determine the Summary Statistics for June
- Deliverable 2: Determine the Summary Statistics for December
- Deliverable 3: A written report for the statistical analysis (README.md)

Upload the following to your surfs_up GitHub repository:

1. The `SurfsUp_Challenge.ipynb` file.
2. The `hawaii.sqlite` file.
3. An updated README.md that has your written analysis

To submit your challenge assignment in Canvas, click Submit, then provide the URL of your surfs_up GitHub repository for grading. Comments are disabled for graded submissions in BootCampSpot. If you have questions about your feedback, please notify your instructional staff or the Student Success Manager. If you would like to resubmit your work for an improved grade, you can use the **Re-Submit Assignment** button to upload new links. You may resubmit up to 3 times for a total of 4 submissions.

IMPORTANT

> Once you receive feedback on your Challenge, make any suggested updates or adjustments to your work. Then, add this week's Challenge to your professional portfolio.

NOTE

> You are allowed to miss up to two Challenge assignments and still earn your certificate. If you complete all Challenge assignments, your lowest two grades will be dropped. If you wish to skip this assignment, click Next, and move on to the next Module.

Module-9 Rubric

Criteria	Ratings					Pts
Deliverable 1: Determine the Summary Statistics for June	**40 to >36.0 pts** **Mastery** ✓There is a working query that retrieves the temperatures from the Measurements table. ✓The temperatures are added to a list. ✓The list of temperatures is converted to a DataFrame. ✓A summary statistics table is generated.	**36 to >32.0 pts** **Approaching Mastery** ✓The query retrieves the temperatures from the correct month. ✓The temperatures are added to a list, but there is a minor error. ✓A summary statistics table is generated.	**32 to >26.0 pts** **Progressing** ✓The query retrieves the temperatures from the correct month, but there is additional data. ✓The temperatures and additional data are added to a list. ✓The temperatures and additional data are converted to a DataFrame. ✓A summary statistics table is generated.	**26 to >0.0 pts** **Emerging** ✓The query retrieves the incorrect temperatures. ✓The incorrect temperatures are added to a list. ✓The list of temperatures and the extra data are converted to a DataFrame. ✓A summary statistics table is generated, but the data is incorrect.	**0 pts** **Incomplete**	40 pts
Deliverable 2: Determine the Summary Statistics for December	**40 to >36.0 pts** **Mastery** ✓There is a working query that retrieves the temperatures from the Measurements table. ✓The temperatures are added to a list. ✓The list of temperatures is converted to a DataFrame. ✓A summary statistics table is generated.	**36 to >32.0 pts** **Approaching Mastery** ✓The query retrieves the temperatures from the correct month. ✓The temperatures are added to a list, but there is a minor error. ✓A summary statistics table is generated.	**32 to >26.0 pts** **Progressing** ✓The query retrieves the temperatures from the correct month, but there is additional data. ✓The temperatures and additional data are added to a list. ✓The temperatures and additional data are converted to a DataFrame. ✓A summary statistics table is generated.	**26 to >0.0 pts** **Emerging** ✓The query retrieves the incorrect temperatures. ✓The incorrect temperatures are added to a list. ✓The list of temperatures and the extra data are converted to a DataFrame. ✓A summary statistics table is generated, but the data is incorrect.	**0 pts** **Incomplete**	40 pts
Deliverable 3: Structure, Organization, and Formatting	**6 to >5.0 pts** **Mastery** The written analysis has ALL of the following: ✓There is a title, and there are multiple sections. ✓Each section has a heading.	**5 to >3.0 pts** **Approaching Mastery** The written analysis has ALL of the following: ✓There is a title, and there are multiple sections. ✓Each section has a heading. ✓There are images, and they are formatted and displayed correctly with one or two	**3 to >2.0 pts** **Progressing** The written analysis has ALL of the following: ✓There is a title, and there are multiple sections. AND ONE of the following: ✓Each section may have a heading. ✓There are images to code, and they are formatted and displayed	**2 to >0.0 pts** **Emerging** The written analysis has ALL of the following: ✓There is a title. ✓There are no headings for each section, but there are three sections.	**0 pts** **Incomplete**	6 pts
Deliverable 3: Analysis	✓There are images, and they are formatted and displayed correctly.					14 pts

Criteria	Ratings					Pts
	14 to >12.0 pts **Mastery** ✓The purpose is well defined. ✓All THREE major points are described. ✓There is a high-level summary of the results and TWO additional queries.	**12 to >10.0 pts** **Approaching Mastery** ✓The purpose is well defined. ✓All THREE major points are described. ✓There is a high-level summary of the results and ONE of TWO additional queries.	**10 to >7.0 pts** **Progressing** ✓The purpose is well defined. ✓TWO of the THREE major points are described. ✓There is a high-level summary of the results and ONE of TWO additional queries.	**7 to >0.0 pts** **Emerging** ✓The purpose is well defined. ✓ONE of the THREE major points is described. ✓There is a high-level summary of the results only.	**0 pts** **Incomplete**	
						Total Points: 100