# Surfs\_Up

![oahu-surfing-experience](https://user-images.githubusercontent.com/85518330/129450037-650f4329-5e88-45d3-8d22-e6a6ca120055.jpg)

## Project Overview

I love Oahu, was there last year for a holiday and want to go back and settle there. What will I do there, you ask? Set up a Surf and Ice Cream shop ofcourse. I think the weather is perfect all year around and toursists and locals alike will Surf first and then cool down from the activity with their favorite Ice Cream flavor. I have a little bit of money but will need some investor to believe in my idea too and help me with some additional funding.

W.Avy my potential investor is excited about my idea and believes it can work not only just in Oahu but also in some surrounding cities as well. A Surf and Ice Cream chain... I can't believe my luck. W.Avy has asked me to do a check on the weather conditions in Oahu through the year just to confirm that its mostly conducive to Surfing and Ice Cream eating. That's easy I say, after all I have been learning data analysis and visualization and honning my data skills precisely for an occasion like this.

## Task Planning

W Avy and his friends my potential investors are not too tech savvy, they are lucky enough to spend much of their time on the beach surfing. I need to make it easy for them to view and understand my analysis. I am also not sure of the Internet speeds in Oahu so I decide to use \*\*SQLite\*\* and \*\*SQL Alchemy\*\* and \*\*Python\*\* to do my analysis, I will then load my results using the \*\*FLASK App\*\* and then W Avy can see and share my analysis with whoever he needs to easily.

## The Actual Analysis

I don't want to confuse W Avy with too much analysis and data, just enough to give him comfort that our choice of Oahu is perfect weather wise for us to set up our Surf and Ice Cream Shop. What should I call my shop I drift off to think. That can wait I quickly decide. I remind myself not to count my chickens before they hatch.

First I decide to pick two points in the year 6 months apart to do my analysis. June the beginning of summer and December seem like perfect choices for my analysis. Afterall they also coincide with the peak tourist seasons for Oahu.

The data I pick spans a period of 7 years from 2010 to 2017. A large enough period to take care of any sudden or unexpected abberations in weather. The data set has two primary tables the \*Measurement table\* and \*Stations table\*

We will use the Measurement table for our analysis as it has the temerature observasions and date details that we want. Next we import certain libraries that we need to run our analysis. We import Pandas as pd and extract from SQLAlchemy amongst a few others

### June Temperatures

We query the SQLite database to get the date and temperature observations, since we need only the results for June we extract the results for the month of June which is the 6th month using the below

\* june\_temps = session.query(Measurement.date, Measurement.tobs).filter(extract('month', Measurement.date) == 6)

Next for easy reading we create a list of the temperatures using the below code

\* june\_temps\_list = [record[1] for record in june\_temps]

Using the list we just created we create a dataframe called june\_temps, we rename the temperature column as June Temp for easy identification

\* june\_temps = pd.DataFrame(june\_temps\_list, columns=["June Temp"])

We run the summary statistics on our dataframe and get the below result

<img width="110" alt="june\_temps" src="https://user-images.githubusercontent.com/85518330/129455078-342e472f-00b7-4dc4-a714-f113d85ae2d6.png">

### December Temperatures

Next repeat the same analysis as above for the month of December

\* dec\_temps = session.query(Measurement.date, Measurement.tobs).filter(extract('month', Measurement.date) == 12)

Next for easy reading we create a list of the temperatures using the below code

\* dec\_temps\_list = [record[1] for record in dec\_temps]

Using the list we just created we create a dataframe called dec\_temps, we rename the temperature column as Dec Temp for easy identification

\* dec\_temps = pd.DataFrame(dec\_temps\_list, columns=["Dec Temp"])

We run the summary statistics on our dataframe and get the below result

<img width="114" alt="dec\_temps" src="https://user-images.githubusercontent.com/85518330/129455204-38f85326-b76b-4310-8256-e7f9db060762.png">

## Inferences

Seeing the summary stats for June and Dec I let out a big sigh of relief. W Avy will be happy to see this.

\* Its surfing times in Oahu for most of the year with average temperatures ranging from 71-75 degrees

\* The low std deviations indicate that the variations in temp for the two the months is not too high

\* One can expect a few chilly days not suitable to Surfing or Ice Cream in December, but hopefully not too many

\* But if we wanted to split hairs then the below are the key differences between the June and Dec statistical summaries

\* The average temps in June are 74.9 and in Dec are 71

\* Dec has a slightly higher std deviation at 3.75 when compared to the June std deviation of 3.25

\* The lowest temperatue in Dec is 56 while the lowest temp in June is 64

\* The highest temps in Dec are 83 while the highest temps in June is 85

## Additional Analysis

Let me also check if rain will play a spoil sport in our plans. Let me give W.Avy and friends an analysis on the precipitation before they ask for it. Precipitation analysis for the same periods tells us the following

First I extract precipitation data from the measurement table and filter it by month for June and Dec using the following query

\* june\_prcp = session.query(Measurement.date, Measurement.prcp).filter(extract('month', Measurement.date) == 6)

\* dec\_prcp = session.query(Measurement.date, Measurement.prcp).filter(extract('month', Measurement.date) == 12)

Next I create a list of precipitations for the month of June and Dec using

\* june\_prcp\_list = [record[1] for record in june\_prcp]

\* dec\_prcp\_list = [record[1] for record in dec\_prcp]

Created a dataframe for the same using

\* june\_prcp = pd.DataFrame(june\_prcp\_list, columns=['June Prcp'])

\* dec\_prcp = pd.DataFrame(dec\_prcp\_list, columns=['Dec Prcp'])

Got the summary statistics for June and Dec

<img width="153" alt="June\_prcp" src="https://user-images.githubusercontent.com/85518330/129462180-437297cb-d42e-4d4e-950c-90864b60f7f0.png">

<img width="123" alt="Dec\_prcp" src="https://user-images.githubusercontent.com/85518330/129462184-3716bbd9-a87c-43e9-b058-b68250656ea9.png">

### Key Findings

\* The precipitation analysis revels that it rains a lot more in Dec as compared to June in Oahu with average prcp of .23 in Dec as compared to .14 in June

\* The maximum prcp in Dec is 6.4 while in June it is 4.4

\* Our descriptive statistics tell us that rain is not going to play spoil sport too often which is quite a relief

# How does the name 'SURF SHACK' sound?