UDACITY

< Back to Data Analyst Nanodegree

Explore Weather Trends

REVIEW

HISTORY

Meets Specifications

Analysis

- **The SQL query used to extract the data is included.**
- **The query runs without error and pulls the intended data.**

Optional I would recommend playing around with `where x like 'string'`. This is a soft match vs an exact match. In some cases where you might not know the exact values of the data, a soft match is more useful.
To Exceed Expectations It is possible to actually get the data in a single query after you get the list of cities you can choose from. It would be could training to learn how to do this. This site helped me a lot as a student. I would recommend attempting to query the entire data in one pass. Here is an example...

```
SELECT city_data.year, city_data.avg_temp AS city_avg_temp,
global_data.avg_temp AS global_avg_temp
FROM city_data
JOIN global_data
ON city_data.year = global_data.year
WHERE city_data.city LIKE 'Bangalore'
```

Moving averages are calculated to be used in the line chart.

Over the Top Amazing demonstration of using a programming language instead of just a spreadsheet. Terrific!

- **A line chart is included in the submission.**
- **The chart and its axes have titles, and there's a clear legend (if applicable).**

- **The student includes four observations about their provided data visualization.**
- **The four observations are accurate.**

One of the best things I learned from an Udacity coach when I took this class pertains to correlation and causation. It is one of the ideas we proposed for students to explore in this project. This is very important when we notice events that seemingly are moving in similar patterns, like global temperature and local temperature. In this data, it is obvious to us since global temps include local temps in its averaging, that we will find correlation which is not necessarily causation. However, in the future of your analytics career, these relationships will not be easy to distinguish. I've put together a sort of helper. I hope it finds you well.

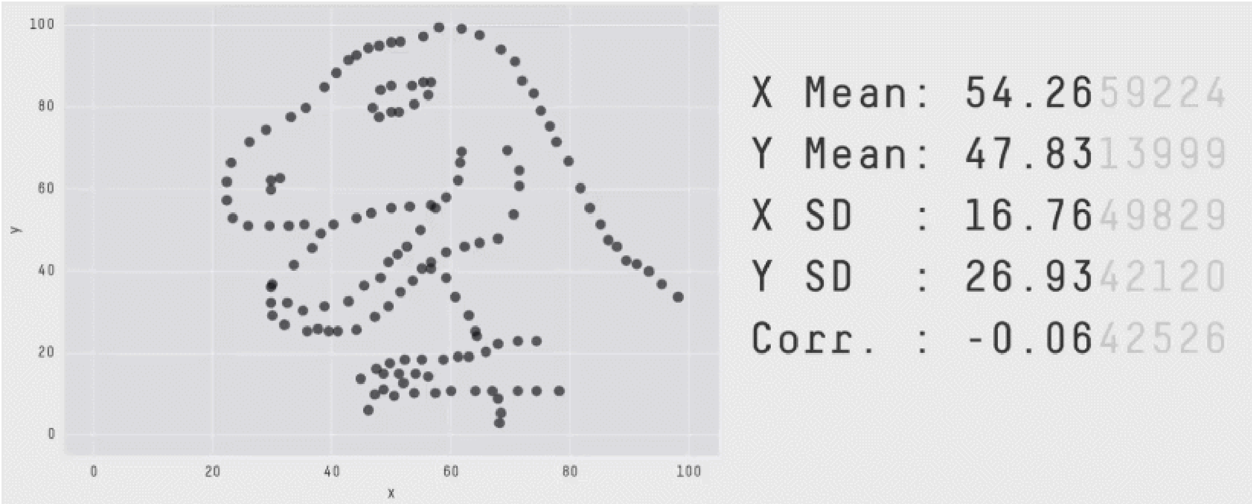
Correlation and Causation

- Generally, you can think of causations as a correlation of 1 (perfect correlation) or -1 (perfect inverse correlation).
- However, not all correlated variables mean causation. To illuminate why, let's talk about the 7 ways we can have correlated variables. For any correlated variables, A and B.
    1. A causes B; (direct causation)
    2. B causes A; (reverse causation)
    3. **A and B are consequences of a common cause, but do not cause each other** *(highlighted for its applicability in our current investigation)*
    4. A and B both causes C, which is (explicitly or implicitly) conditioned on. If A and B cause C, why do A and B have to be correlated?;
    5. A causes B and B causes A (bidirectional or cyclic causation);
    6. A causes C which causes B (indirect causation);
    7. There is no connection between A and B; the correlation is a coincidence.
- So as you can see, in order to determine *causation*, we must eliminate the possibilities of just correlation by asking if any of those causations are possible.
    - Then we must be able to prove empirically the causation.

Here is what you should ask...
Can you assess which of the above situations of causation we are witnessing?
How can we prove this?

Correlation does not Imply Causation
Funny Spurious Correlations

All the following graphs have the same statistics including correlation.



Citation: https://www.autodeskresearch.com/publications/samestats

DOWNLOAD PROJECT

RETURN TO PATH