Introduction

Saguaro population/collection at DBG

History of Saguaro Inventory

Research questions

Conclusion

Summarize clean-up

Summarize analysis

Discuss the problems encountered during project

Git/GitHub merge conflicts

Steve’s “negative feedback loop” 😊

In order to count the number of days in each timespan with nighttime low >+ 90, we required 8 years of daily, historical weather. Steve’s comments on lack of availability of free historical daily weather API’s.

Pd.DateTimeIndex function would not operate on our datetime format so Steve had to learn regex to parse the year out of each datetime object using the space.

Chained indexing / the UnderstandingSettingWithCopyWarning error

Very dirty dataset:

Duplicates

Missing data – removed saguaros missing data

3-dimensions – broke the dataset into two to reduce to two dimensions

Data in inconsistent units that had to be standardized

Two timespans of different length. Had to standardize before we could compare.

Data that needed to be parsed – regex!

Future work

This analysis was based on the first two timespans of a study that will likely go for decades.

value1.loc[value1['Total Population'] == '\*', 'Total Population'] = 4

C:\Users\vsnix\anaconda3\envs\PythonData\lib\site-packages\ipykernel\_launcher.py:1: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: <https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy>

"""Entry point for launching an IPython kernel.

df.loc[df[“Measurement Units”] == “inchs”, “Measurement”] = df[“Measurement Units”]/12

is there a way to select all the inches rows and calculate their feet equivalent into an intermediate variable instead, then load it back into the dataframe?