Cannabis Data Science – Meetup 2/10/2021

Hello everyone, my name is Keegan and its awesome to meet each of you. This group is for people who want to apply data science to cannabis data. My background is in economics and lab work, so I have an interesting perspective and abundant interest in the subject, so, I wanted to host a group. I can share some of my work, but this group can go where the group decides, so, please feel free to speak and share. And we’ll take it from there.

What does a data scientist do?

A senior analyst may spend the bulk of their time cleaning data to be used in prediction models. The data is presented in a final report in a simple format. The datasets may be in the thousands to in the billions.

1. The first step is to find or build a dictionary of the data.
2. Next, you need to import all of the data so that you can work with it.
3. Now, you will need to begin to clean the data. This requires an understanding of the structure of the data.
4. Once the data is defined, read, and formatted, then you can begin analysis. This is typically one of the faster stages if data is cleaned and formatted well. Data is then analyzed with mathematical and statistical functions, such as regression analysis, to make comparisons and gain insights.

The first rule of data science: look at the data. Next, you need to formulate a research question. Then you can begin to think about how to clean the data and what variables to keep.

Data scientists may use timeseries regressions. You can find interesting points in time and compare the series from before and after the event to try to identify any structural breaks.

You can control for region, such as urban or rural, and other geographical variables.

It is often incredibly fruitful to combine multiple datasets.

Resources

WSLCB December 2020 Data

<https://lcb.app.box.com/s/fnku9nr22dhx04f6o646xv6ad6fswfy9?page=1>

Cannabis Genome Data

<https://www.kaggle.com/paultimothymooney/how-to-query-the-1000-cannabis-genomes-project>

https://hackersandslackers.com/creating-django-views/