CONCEPT
Create a program that allows users to input a city and find the most
"instagrammable" spots in the area. The program will scrape Instagram for
the most liked posts in a city and filter out influencers and other external
factors in order to return the best specific locations to take an Instagram
photo.

DATA COLLECTION
https://github.com/rarcega/instagram-scraper
We plan to use the above Instagram scraper and its --location flag to find
the top posts in a city, as well as its –latest tag to find the most recent posts
in a city. We'll then use the --user tag on the posts we've found to seed
more post locations in the city to analyze.

DATA CLEANING

RECENT DEVELOPMENTS
Cloud set-up
Sample data

PIPELINE

Choose broad location (Ex. NYC)

Scrape top posts

Choose top 50 sub-locations

We only want to look at posts that have had adequate time to accumulate likes

Scrape top posts

Scrape recent posts

Filter out posts made within 24 hours

(user, time, location, number of likes)

Scrape posters' profiles

Store post data

Filter out external factors - make sure the location is the defining factor, not the person

Seed new locations

Compare ratio of likes of original post to historic posts

Mark as influencer based on number of followers