Assessing Bee Health by Subspecies

Kaila Kay
Flatiron Immersive Data Science: Module 4

Data Set: Kaggle

Slide Formatting:
Farmers Market Coalition

TABLE OF CONTENTS

01
Overview
Description of the project and deliverables

02
Topic Analysis
Honey bees in areas of the United States

03
Data Analysis
Exploring datasets to investigate pre-determined questions

04
Prediction Modelling
Bringing it all together

05
Future Work
Where I would like to see this project go in the future

06
Wrap Up
Conclusions and Gratitude

01

Overview

Investigating subspecies of **Honey Bees** by health, region, and pollen presence

Column Descriptions	
File	File name in bee_imgs folder
Date	Date of video capture
Time	Time of day of video capture (military time)
Location	Location (city, state, country)

	file	date	time	zip code	subspecies	health	pollen_carrying	caste	City	State	Country	is_healthy
2916	010_542.png	8/19/18	12:46	50315	Carniolan honey bee	healthy	False	worker	Des Moines	IA	USA	1
2038	016_270.png	8/20/18	11:55	95124	-1	healthy	False	worker	San Jose	CA	USA	1
3830	032_129.png	8/21/18	9:00	50315	Russian honey bee	healthy	False	worker	Des Moines	IA	USA	1
885	030_374.png	8/17/18	17:50	77511	Italian honey bee	ant problems	False	worker	Alvin	TX	USA	0
4353	038_110.png	8/18/18	12:30	30607	1 Mixed local stock 2	Varroa, Small Hive Beetles	False	worker	Athens	GA	USA	0

Zip Code	Zip Code to numerically describe loaction
Subspecies	Subspecies of Apis mellifera species
Health	Health description of bee
Pollen_Carrying	Presence of pollen on the bee's legs
Caste	Worker, Drone, or Queen bee

Count: 973

Polk County; IA

Cheshire County; NH

Count: 92

Santa Clara County; CA

Count: 2,149

Count: 1,051

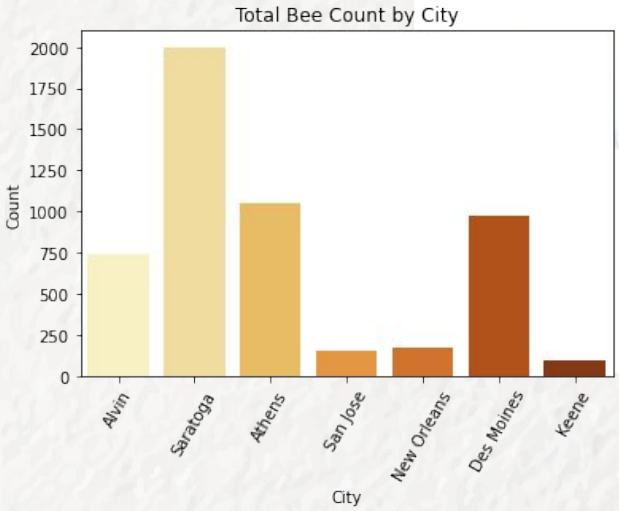
Clarke County; GA

Count: 170

Orleans Parish; LA

Brazoria County; TX

Count: 737



02

Topic Analysis

Investigating subspecies of **Honey Bees** by health, region, and pollen presence

HONEY BEE HEALTH BY NUMBERS

- **33** percent of our global food production relies on pollinators.
- **90** U.S. crops pollinated by commercial honey bees each year.
- **100+** Almond Board-funded honey bee health research projects to date.
- **1995:** The year we created the Honey Bee Health Taskforce to specifically fund research to improve bee health.

Source: Food and Agriculture Organization of the United Nations, 2017.

Image

03

Data Analysis

Investigating subspecies of **Honey Bees** by health, region, and pollen presence

Data: Farmers Market

Engineered

City, State columns

Dropped

Location

Engineered

Counties based off zip codes for mapping

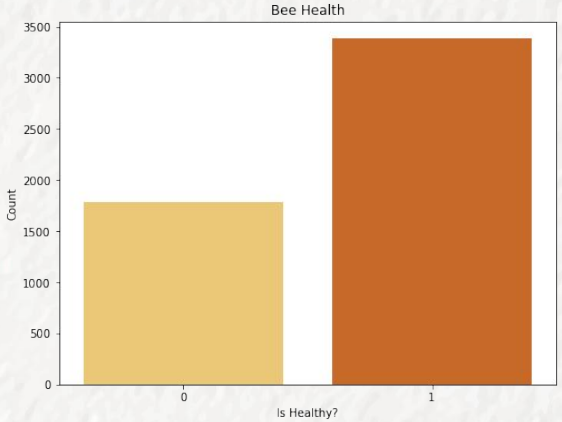
Connected

File column values to respective images

Engineered

Binary encoding on bee health y/n

Bees Carrying Pollen



Bees Carrying Pollen

0.35%

18 Bees

Saratoga: 6
San Jose: 11

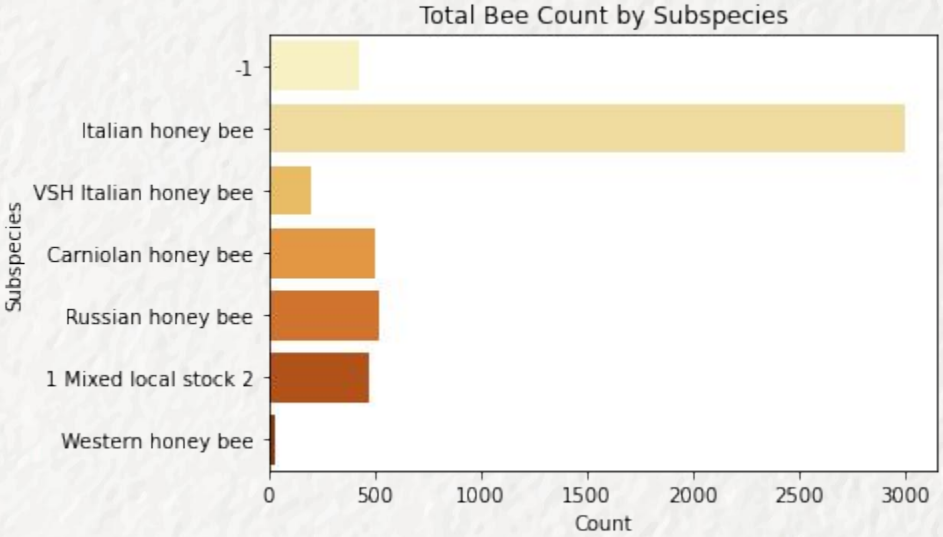
Subspecies

Italian Honey Bee: 6
Undetermined Species: 12

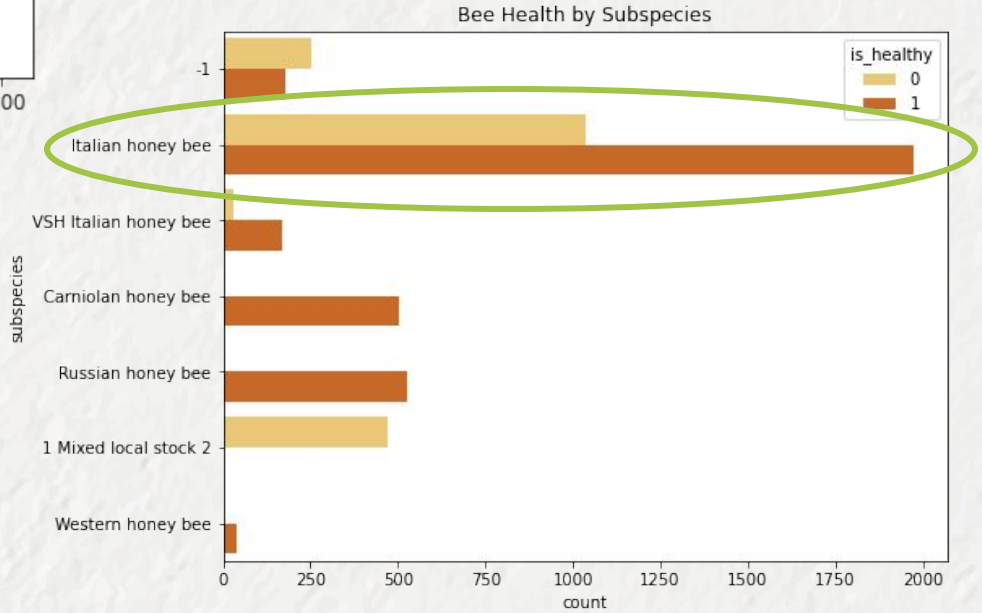
Facts

With the exception of 1 unhealthy bee in Georgia; all pollen carrying bees were recorded in California.

More than 5,000 bees were recorded as non-pollen carrying



The Italian Honey Bee has the highest number of instances in the dataset: **3008**

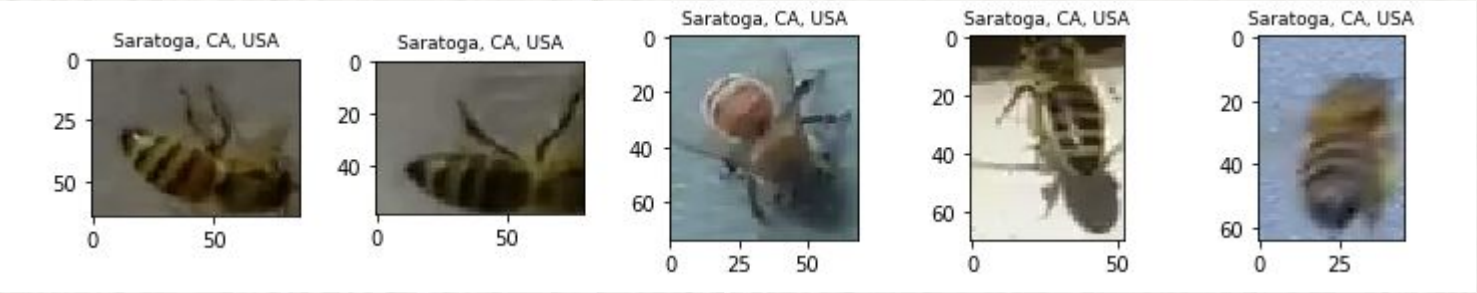


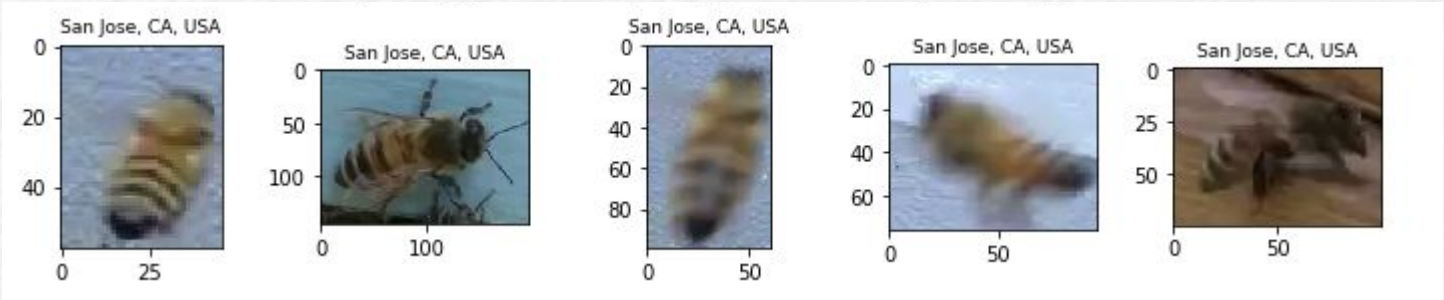
The Italian Honey Bee is split roughly **⅓** unhealthy and **⅔** healthy

EDA

Investigating Italian Honey Bees & California

California Bees:





04

Prediction Modelling

Predicting Bee Subspecies from Images

Accuracy & Precision

Accuracy

89.57%

Precision

Macro Avg: 0.89

Recall

Macro Avg: 0.88

Loss

26.55%

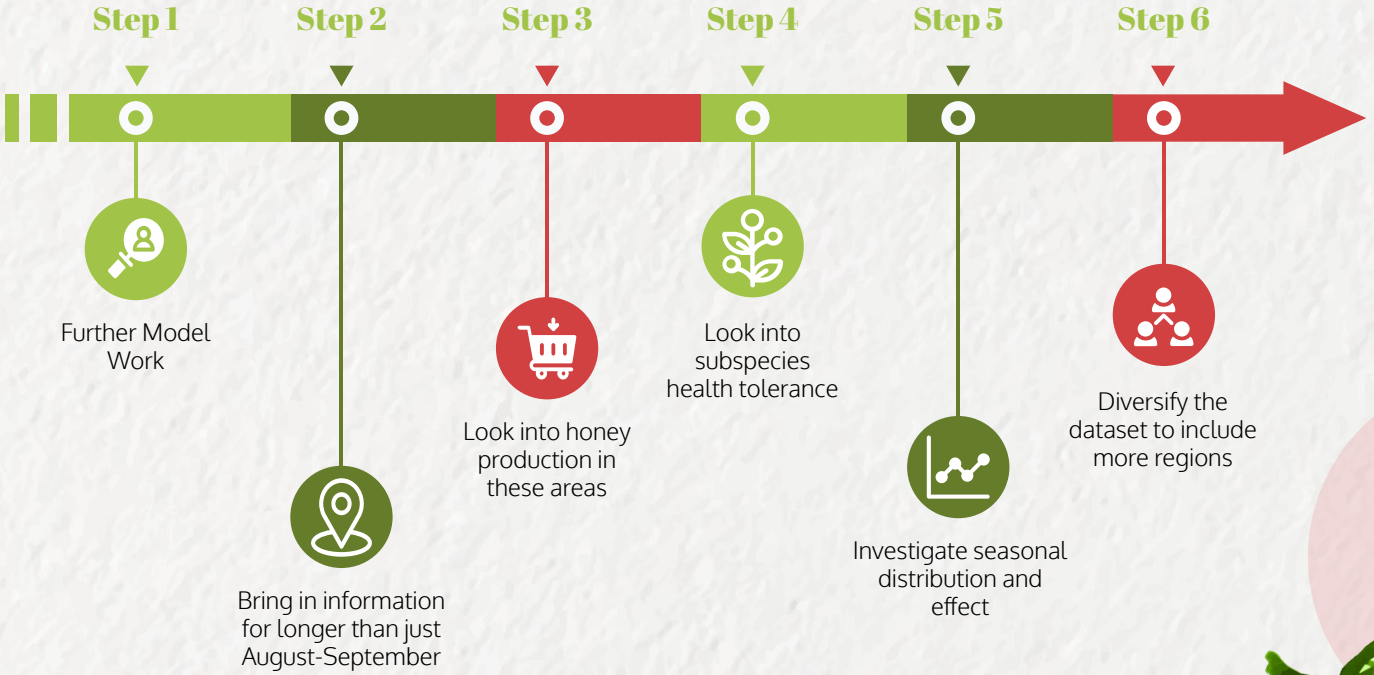
Goals

Find more parameters for training in an effort to increase accuracy and precision

Find more instances of other bee subspecies

Find more data from different times & regions

05 Future Work



06

Conclusion & Wrap Up

Investigating subspecies of **Honey Bees** by health, region, and pollen presence

Thank You!!!