Pokemon Machine Learning

Final Presentation

Brian Gerrard - Linkedin/Github
Jennifer Cheng | LinkedIn/Github
Tatiana Kalainoff | LinkedIn/Github

Agenda

Fill in retroactively

Meet Team Sprocket

Brian Gerrard

Github Link
LinkedIn Link

Jennifer Cheng

Github Link
LinkedIn Link

Tatiana Kalainoff

Github Link
LinkedIn Link

We are students of the
Columbia University | Engineering Data Analytics and Visualization Certification Program
Class of August 2021

We're showcasing a data-led retrospective on 7 Generations of Pokemon

Happy Birthday, Pokemon!

Why we love this project

Pokemon has had a cultural resurgence in recent years, all leading to the 25th Anniversary of Pokemon in October 2021.

This franchise is a cultural phenomenon, and, In 2021, it is celebrating a milestone anniversary while consumers are yearning for nostalgia and simple pleasures.

We're taking a look at how the original Pokedex has evolved across 7 Generations. Pun intended.

The Data

We're using a robust dataset that contains a catalogue of Pokemon that spans the first 7 Generations.

Some stats we're able to analyze are outlined here ➔

Data Column	Description
Pokedex Data	A catalogue of all detailed Pokemon species and features
Base Stats Data	Each individual value that determines a Pokemon's skills, strengths, weaknessess and overall abilities
Training Data	Quantifed efforts for leveling up, upskilling and evolving your pokemon
Breeding Data	Detailed data on which types of Pokemon species can breed and what they evolve into
Type Defenses Data	Data that determines a Pokemon's resistance against physical attacks

Learn more about Pokemon stats

Questions We'll Answer

Not yet committed to the things we're building

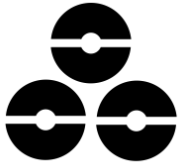
1. Exploration of the characters. Show the distribution of _____ with each generation:
- Types
- Gender
- Abilities
- How to visualize these changes over time?
- Potential Precedent: The Evolution of the US Census
- Potential Precedent: Scientific Proof that Americans are Completely Addicted to Trucks

2. Collecting them all. How does gameplay change with each generation? Compare ____ across generations:
- Catch-rates
- Egg-cycles
- Base Experience
- How to visualize?

3. Fun Facts!
- How many Pokemon do not have evolutions? -- What is their species make-up?
- Who are the top 10% unfriendliest Pokemon?
- Who are the top 10% friendliest Pokemon? -- Are there any trends here? Gender/Type/Species --* Any other features at all - the list is endless

4. What's your Pokemon's aura?
- Shirley Wu - Film Flowers inspo?

5. Interactive Sandbox
- Build your own pokemon? Enter base stats and use machine learning to categorize your new pokemon by type -- would need to know: --- Are there correlations between base stats and the primary type of pokemon? If not, what correlations can we pull? maybe type vs these base stats. example



Coding Techniques

Match each project with what type of coding/skills are used - can be filled retroactively

Outline format?:

‘Name of Section’ (e.g., pokemon builder)
Synopsis of process (e.g., Use unsupervised machine learning to…)
Languages Used (Python, MongoDB, etc.)

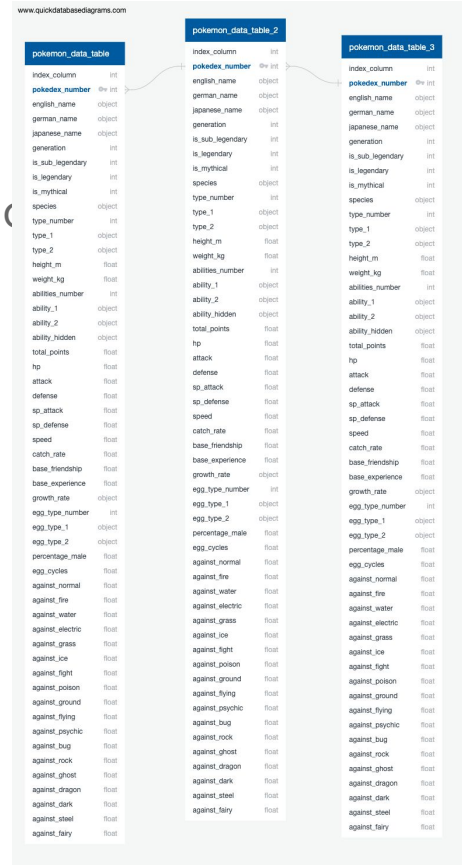
Data Exploration

Description of the data exploration phase of the project

Pokemon ERD

Link to ERD in GitHub -

https://github.com/jwc324/Pokemon_Machine_Learning/blo mo
n_ERD.png



We probably need some sort of data exploration output?

Data Analysis

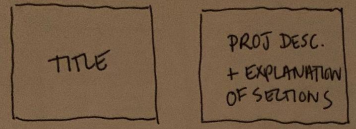
✓ Description of the analysis phase of the project

We probably need some sort of data analysis output?

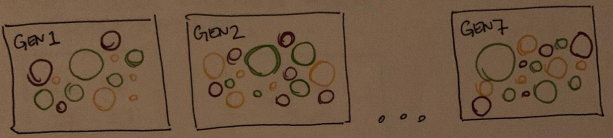
Storyboard

✓ Description of the analysis phase of the project

INTRO



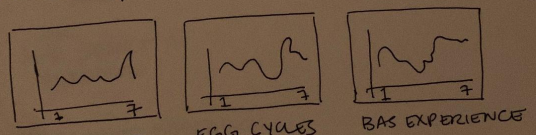
EXPLORATION OF CHARACTERS
[ANIMATED BUBBLE CHARTS]



- TYPES
- GENDER
- ABILITIES

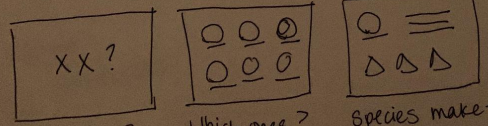
} ability to filter on + off

COLLECTING THEM ALL



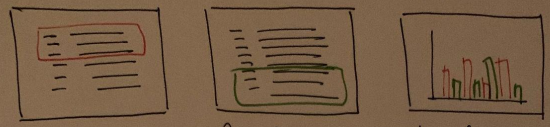
CATCH RATES | EGG CYCLES | BAS EXPERIENCE

FUN FACTS – POKEMON W/O EVOLUTIONS



How many? | Which ones? | Species make-up of one.

FUN FACTS – 10% Unfriendliest / Friendliest



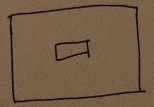
Unfriendliest | friendliest | trends

WHAT'S YOUR POKEMON AURA?
[DATA BADGES]

INPUT:
- TYPE
- GENDER
- ABILITY
- FRIENDLINESS

OR
JUST SELECT A POKEMON

INTERACTIVE SANDBOX



INPUT:
- BASE STATS