App Investment Strategy

**January 2020**

By: Amanda Gaeta

Agenda

- Introduction

- Methodology

- Findings

- Next Steps

Introduction

- This year the firm has been revisiting our tech investment strategy to focus on apps

- Android is the leading mobile operating system worldwide with an 87% percent market share*

- Built a model to predict whether an app will be positively rated at a 4.5+ rating out of 5 on the Google Play Store to help guide our future investments

* Source

Methodology
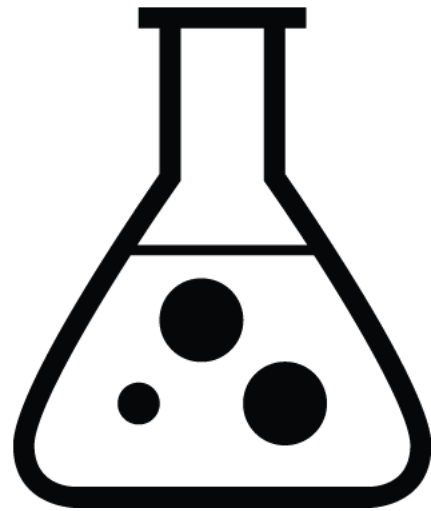
The Data

- 2019 Google Play Store data* representing approximately 10,000 apps

App   **Rating**   **Size**   **Category**

Type   **Genre**   **Reviews**   **Installs**

- Model strictly incorporates **controllable features** so when we understand what drives success, we can edit our approach

The Goal

- Model is built to focus on a metric called **recall** that answers the question:

> **Of all the apps that were truly positively rated, how many did this model identify?**

- A recall focus also means that the model will **avoid False Negatives**. For the firm, this means avoiding losing out on investing in an app that is predicted to be positively rated.

Findings

Model Results

4 Model types tested

14,000+ iterations

56-58% were highest recall scores

Installs, Reviews, and Size of App were most important features

...but this model is not in a final versioning to be utilized in day-to-day operations

Note: goal for recall scores is higher than 50%

Recommendations

- Consistency in important features across models can be generally applied to say that marketing, promotion, and adoption of the app is foundationally important

Along with the technical aptitude for efficient app development, the firm would want CEOs with strong marketing background, a network with these skills, and/or teams that have this expertise.

Next Steps

Next Steps

- Internal operations
  - Applications for funding and questions for in-person pitching
  - Selecting company representatives in later interview rounds
  - Targeting organizations for VC investment.

- Additional Data Analysis and Modeling
  - Cleanse and iterate upon the current data
  - Fold in more Google Play Store data
  - Focus on maintaining or increasing recall and additional feature engineering (i.e. age) to create post-funding milestones for the app companies to achieve
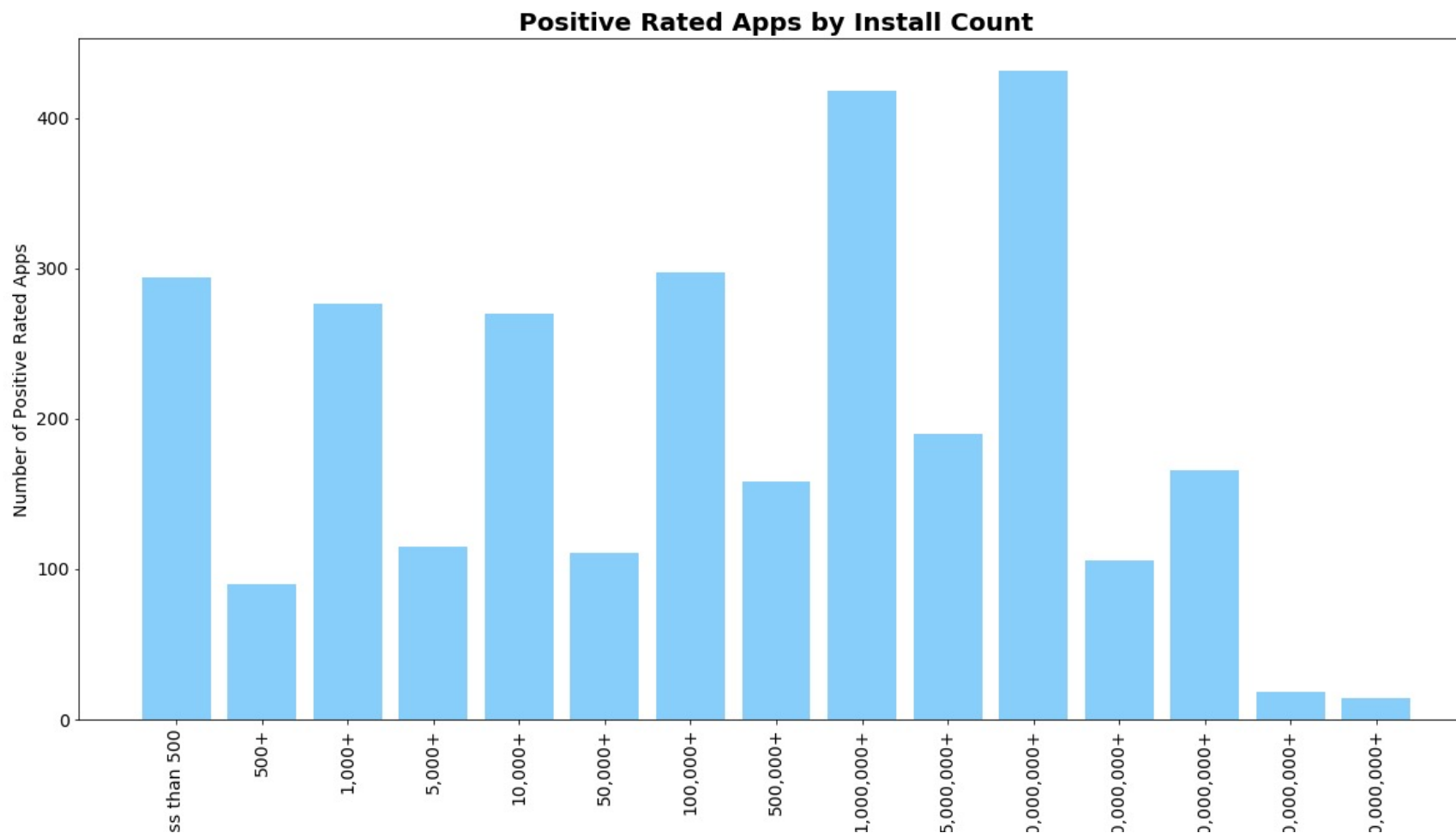
Thank you!

Appendix

Goal: Recall Detail

Model is built to focus on **identifying True Positives** and **minimize False Negatives**. This means minimizing missed opportunities to invest in positively rated apps.

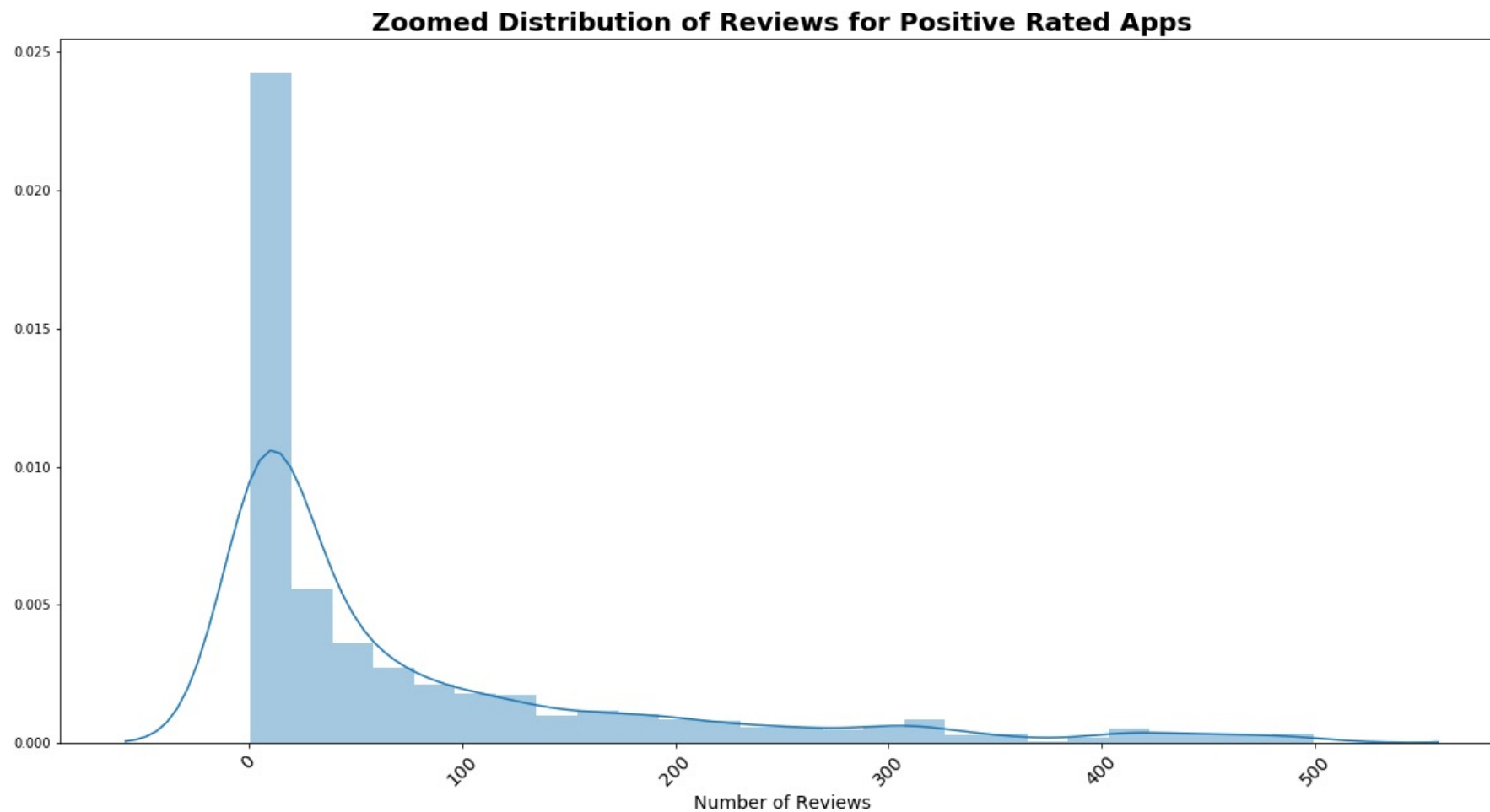|	**1**	**0**
1	**True Positive** An app is predicted to be positively rated and it is	**False Positive** An app is predicted to be positively rated and it isn't
0	**False Negative** An app is predicted to not be positively rated and it is.	**True Negative** An app is predicted to not positively rated and it isn't

Feature Importance: Installs

Feature Importance: Reviews

Zoomed Distribution of Reviews for Positive Rated Apps

Feature Importance: Size

Distribution of Size in Kilobytes (kb) for Positive Rated Apps

Size in Kilobytes (kb)