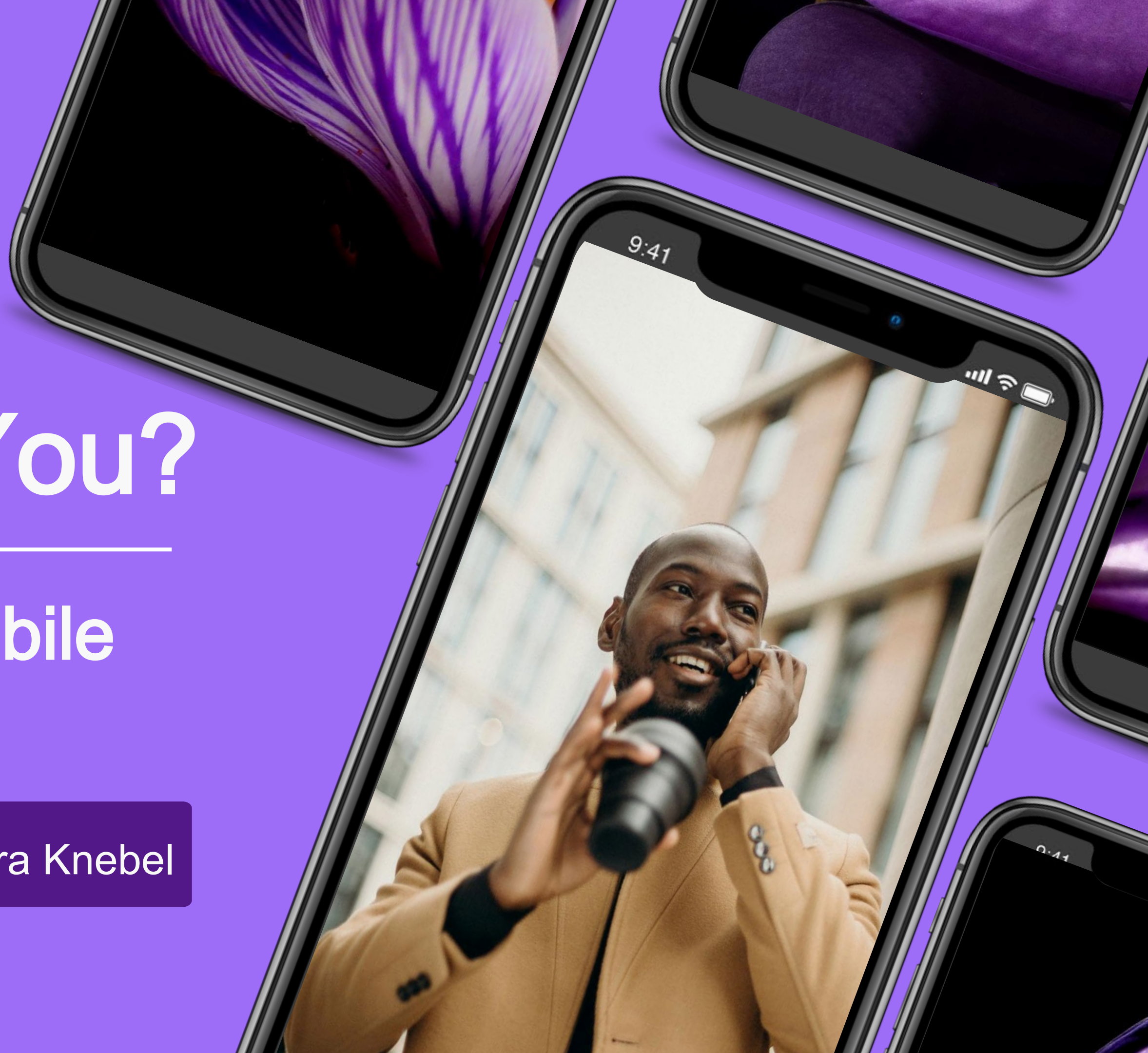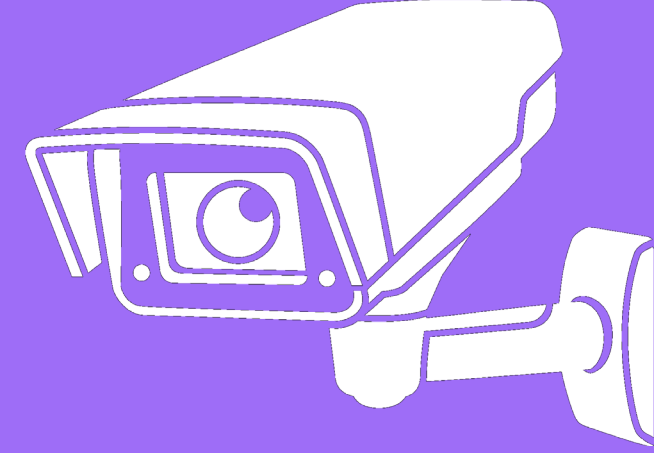# Who's Watching You?

An Analysis of Mobile

App Trackers

Sarah Mayner, Mit Shah, Laura Knebel

Roadmap

Project Overview

Data Summary

Model and Analysis

Recommendations

Motivation

- Number of third party trackers in apps rising

- Apps track online + offline activity

- Identify categories of apps that collect an unnecessary amount of data on users

- Allow users to make more informed app choices

- Inform future policy decisions

Defining Terms

Permissions:
Rights an app asks for to access your phone

14 **permissions**

We have found the following permissions in the application:

⚠ ACCESS_COARSE_LOCATION
access approximate location (network-based)

⚠ ACCESS_FINE_LOCATION
access precise location (GPS and network-based)

ACCESS_NETWORK_STATE
view network connections

ACCESS_WIFI_STATE
view Wi-Fi connections

Trackers:
Software that gathers info on usage/phone

Dangerous: App wants data and resources that involves the user's private information

Signature: System grants these app permissions at install time, but only when the app that attempts to use a permission is signed by the same certificate as the app that defines the permission

Normal: No risk to user's privacy or the operation of other apps

Benign: Some permissions were missing a permission level mapping (dangerous, signature, or normal). Upon inspection, they did not appear to be dangerous; we labeled them benign to increase dataset

Dataset

2.5M

Android apps

~40K

w/ permissions data

3K

w/ ≥1 categorized
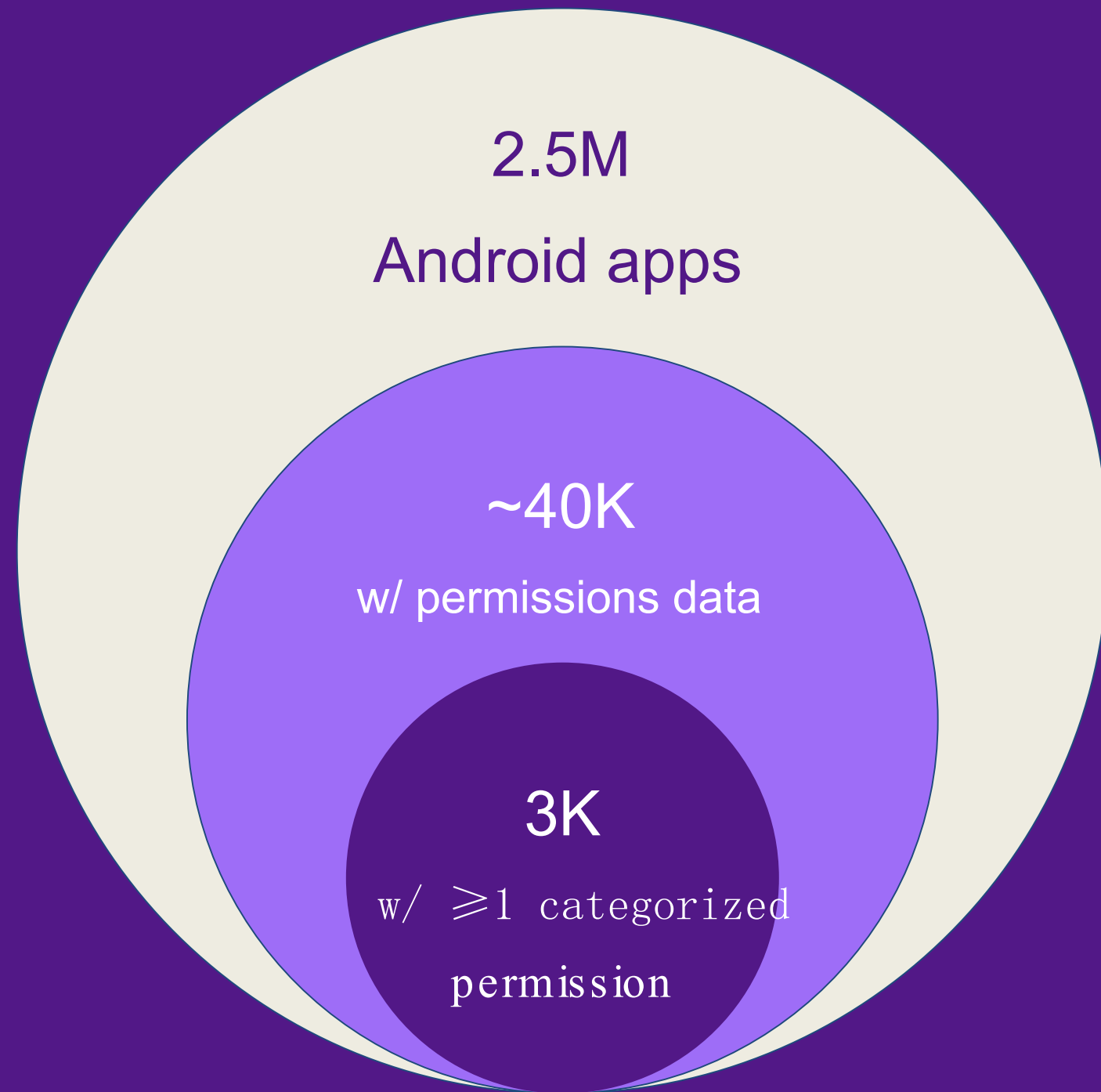
permission

DATA CLEANING PROCESS
- Data source included consumer facing play store apps, malware apps and third party apps with 1400+ permissions
  - Segmented initial dataset of 100k play store apps
  - Dropped 60k rows of apps with duplicate details
- 170+ permission mapping taken from Android Developer Guide
  - 120+ permission categorized as normal, signature , or dangerous
  - Uncategorized permissions marked as benign
- Apps retained with at least one mapped permission
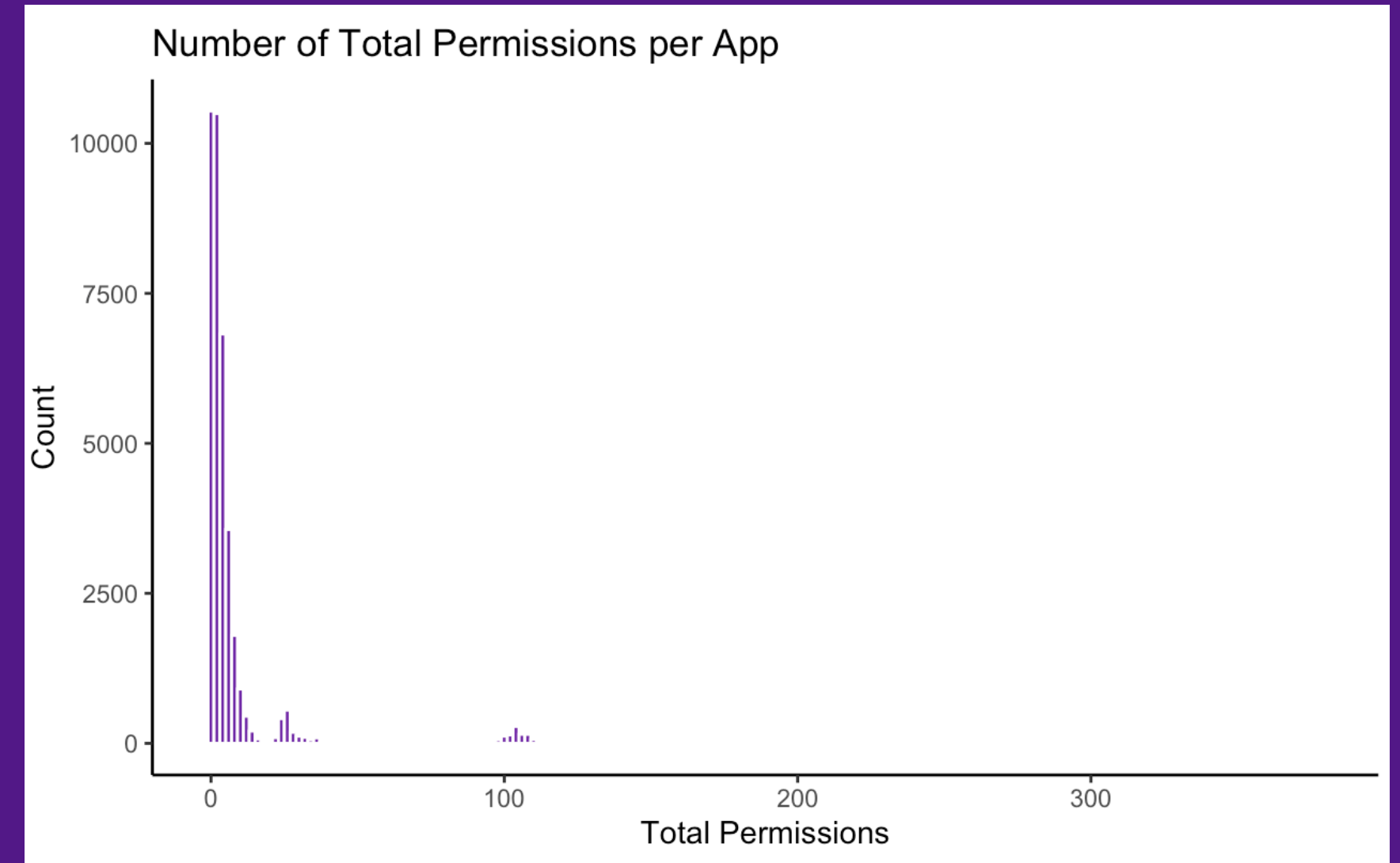 - Apps with none could have non-mapped dangerous permissions

VARIABLES OF INTEREST
- Number of Total Permissions
- Number of "Dangerous" Permissions
- Category of App

Data Sources:
Mahindru, Arvind (2020), "Android permissions dataset, Android Malware and benign Application Data set (consist of permissions and API calls)", Mendeley Data, V3, doi: 10.17632/b4mxg7ydb7.3
"Android Developer Guide." https://developer.android.com/reference/android/Manifest.permission

Project Overview   **Data Summary**   Model and Analysis   Recommendations

Data Exploration: Totals

Number of Apps in Each Category


Number of Total Permissions per App
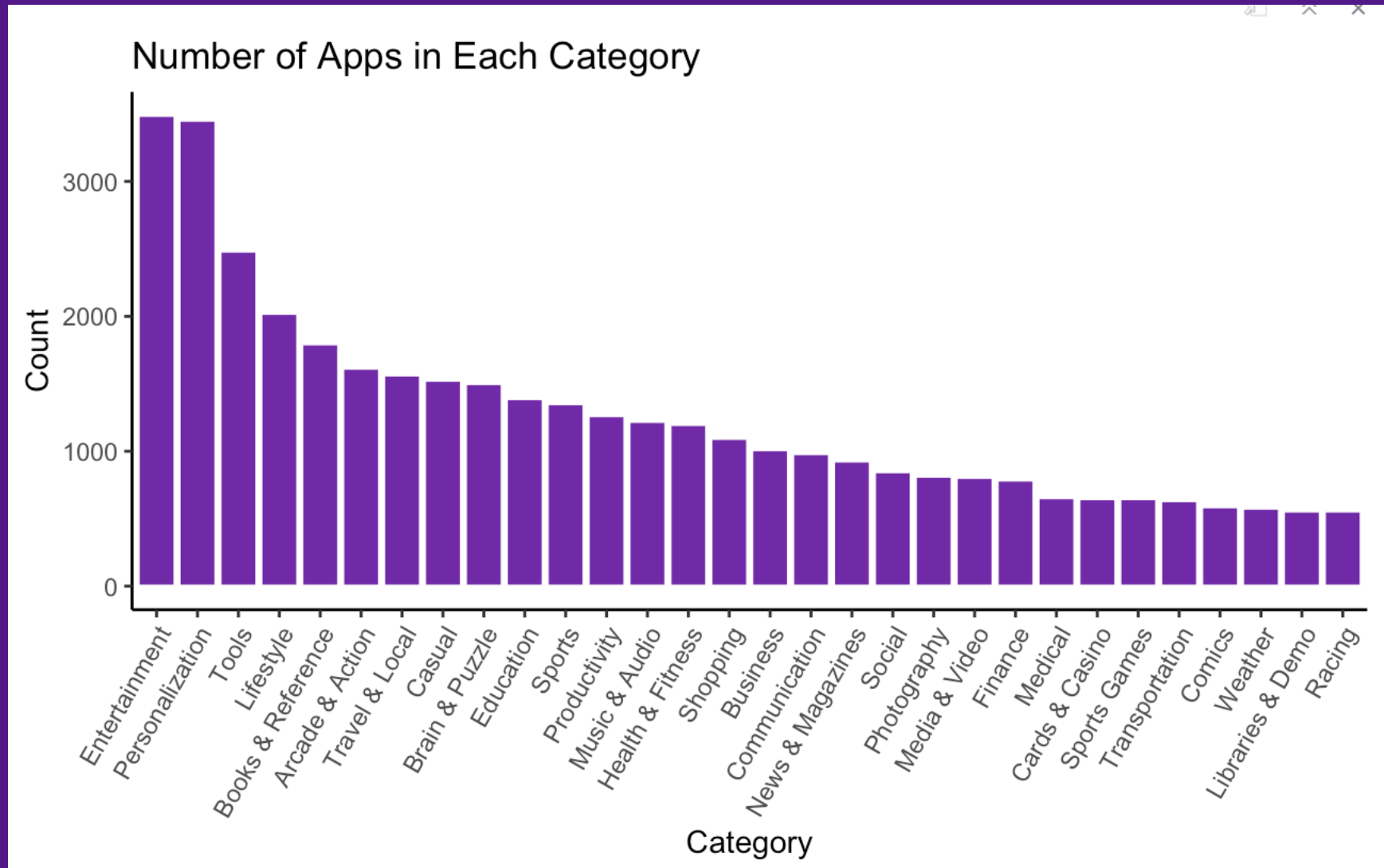
Top 5: entertainment, personalization, tools, lifestyle, books/ref

Permissions may be more important for apps used daily

More app competition within a category means users can select apps with more responsible permissions

Most apps have only a few permissions enabled (75% have ≤ 20)

However, some have an egregiously large number (max = 378)

Data Exploration: By Category

App Privacy Protection Level by Category


Total Number of Permissions by App Category
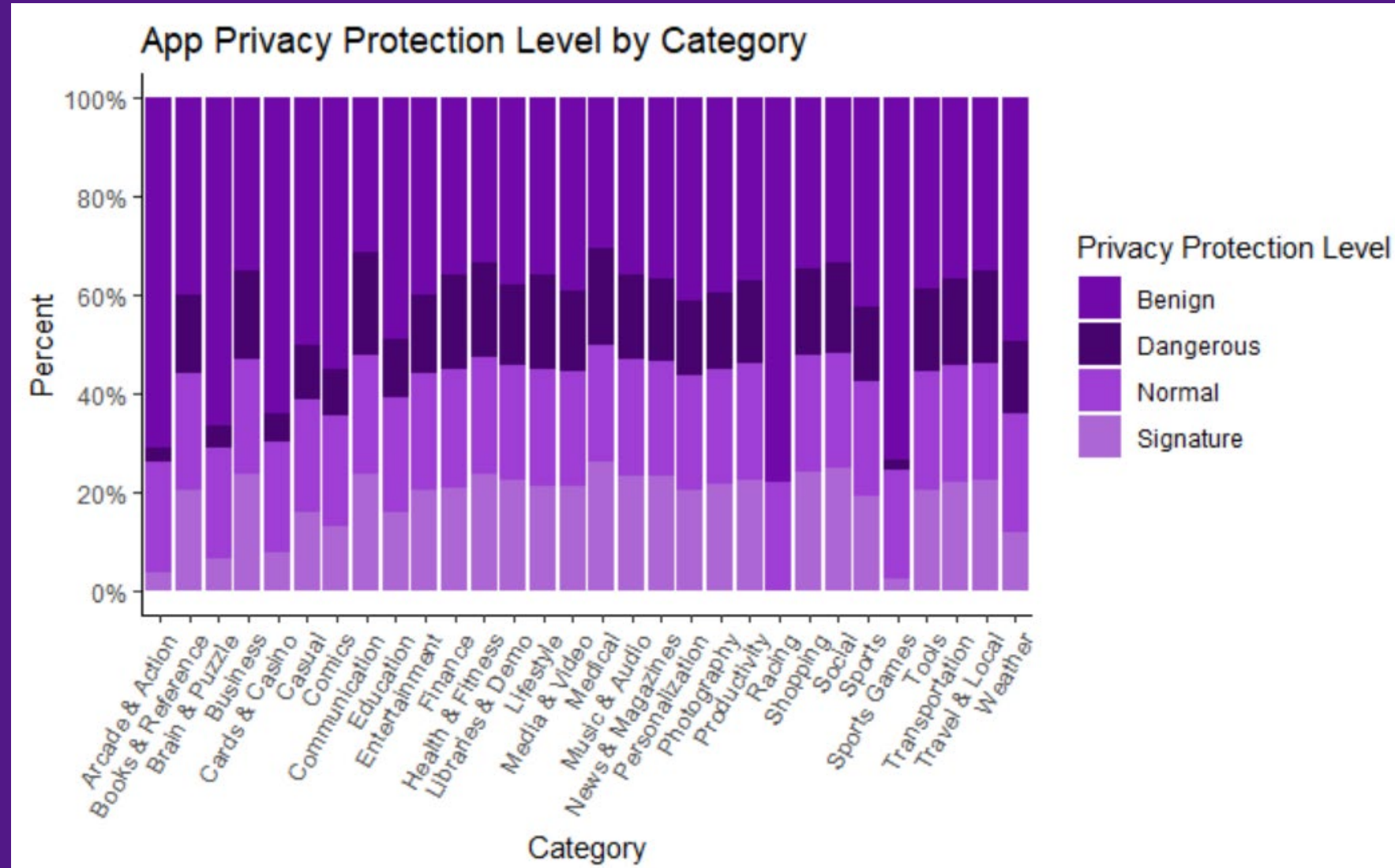
The proportion of permissions in each category that fall into each privacy level varies, with some categories (like Sports Games) having more benign/normal rather than dangerous.
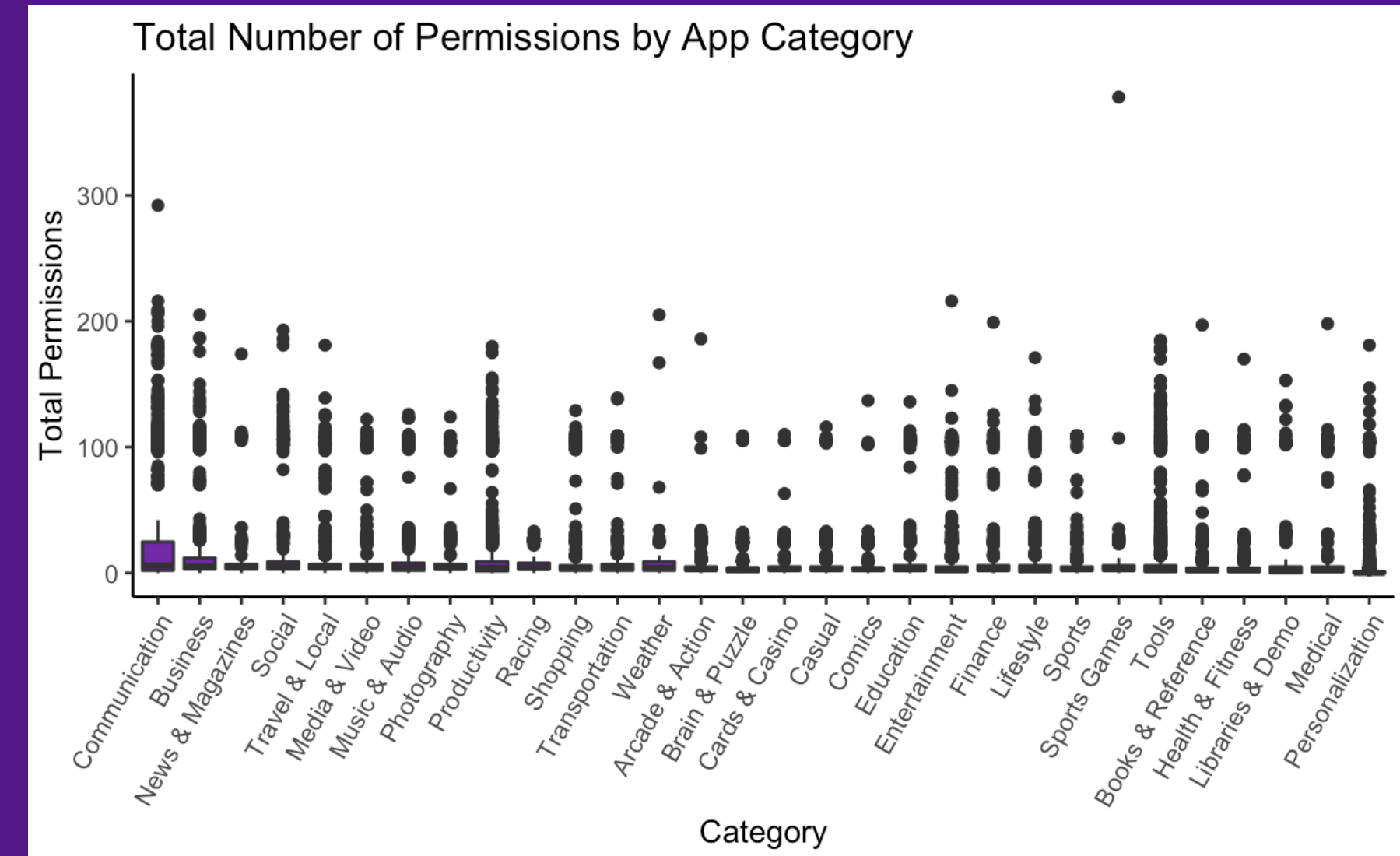
Lots of variation within each category; all have outliers
Highest: communications, business, news, social
Lowest: personalization, medical, libraries, health

Data Exploration: DANGER

Number of Dangerous Permissions by App Category

Number of Total vs. Dangerous Permissions by App Category
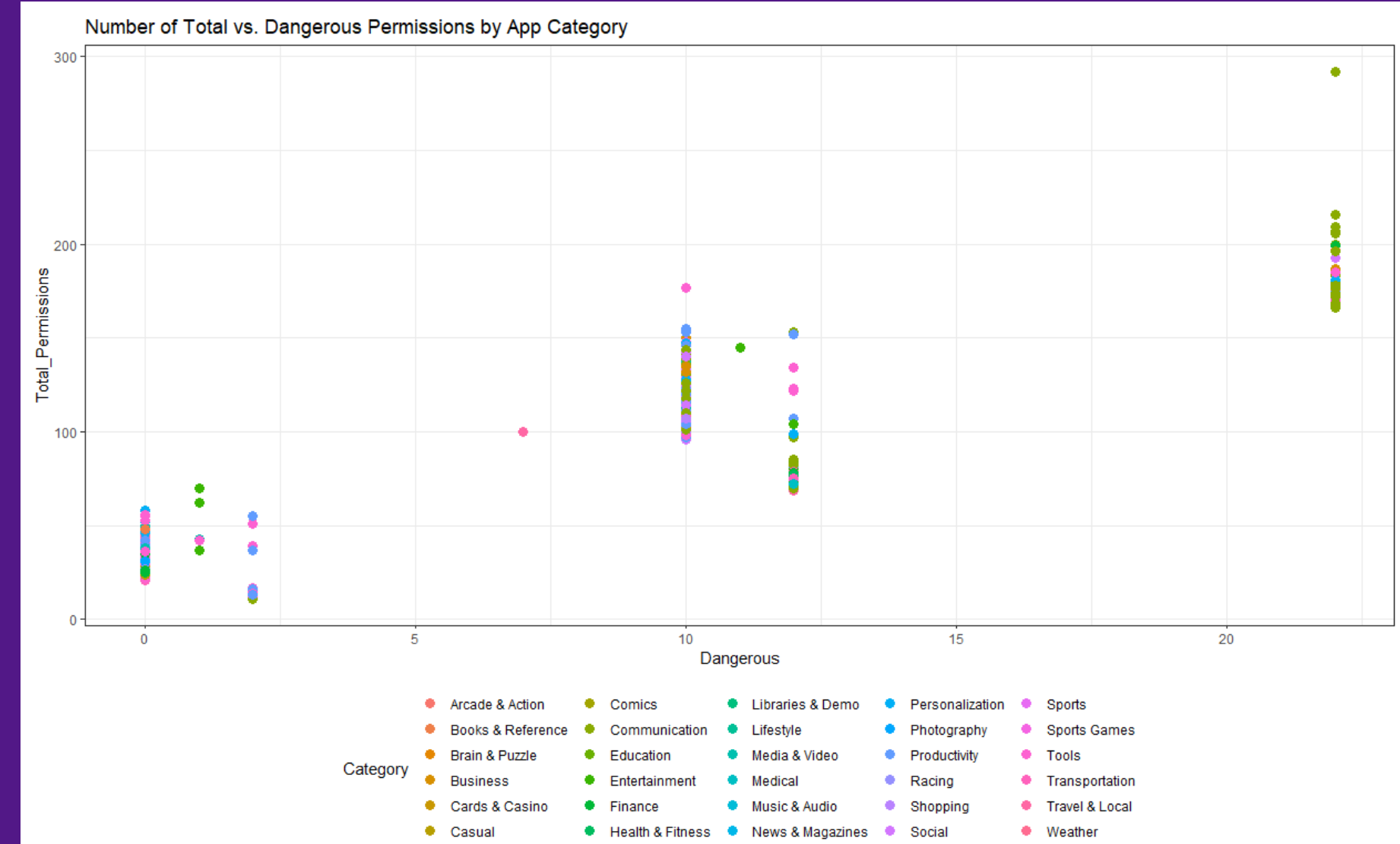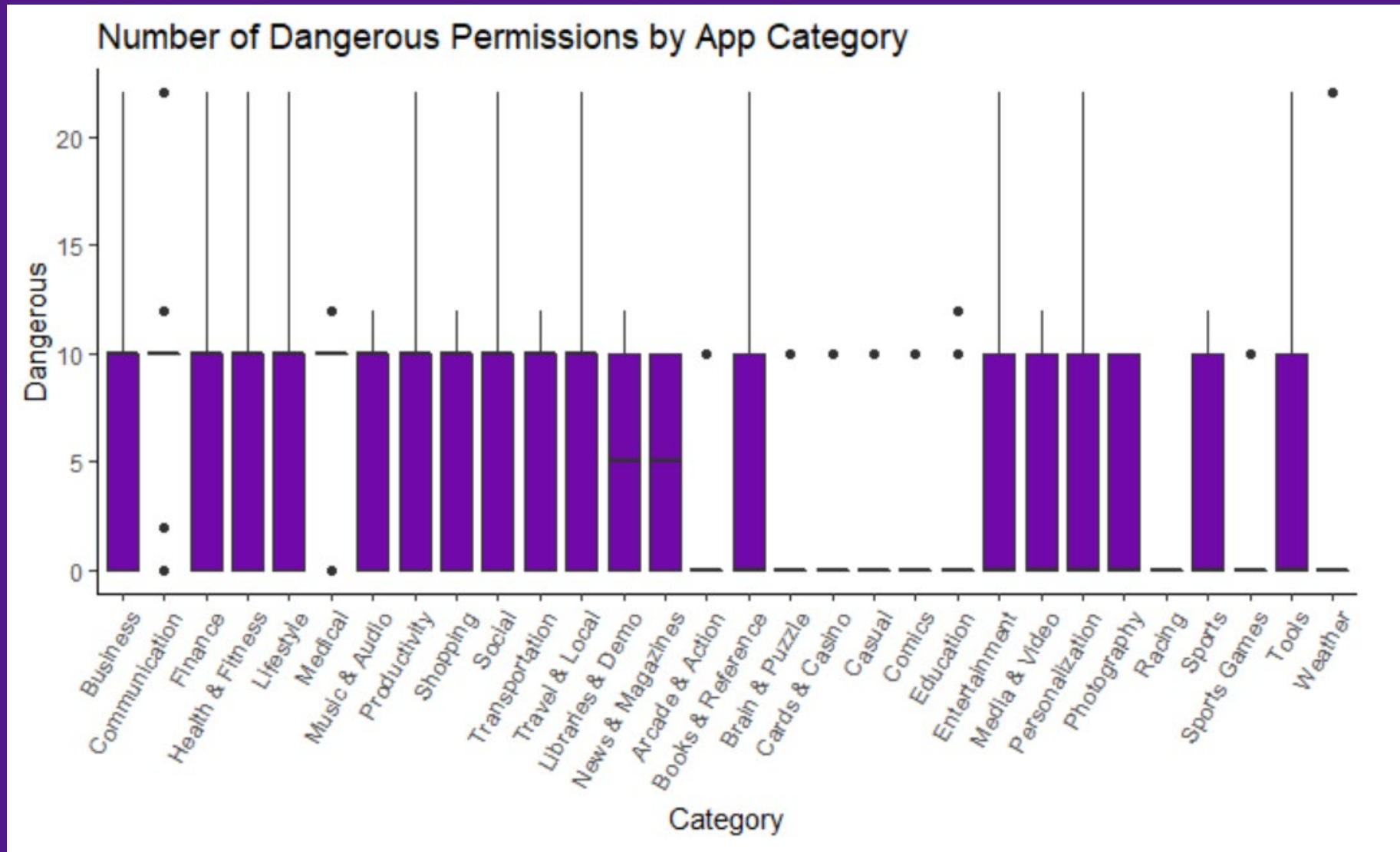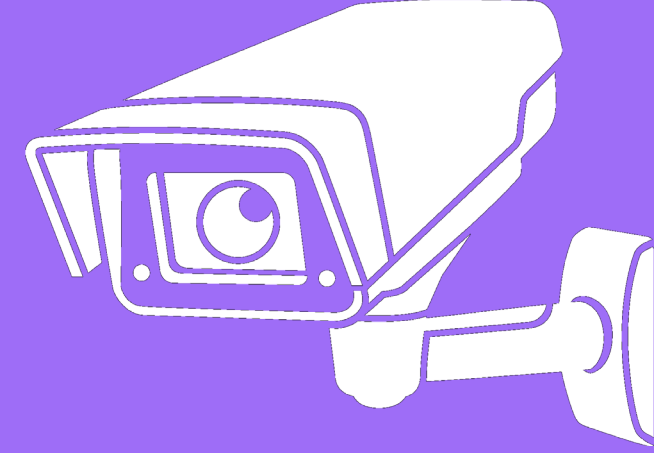
Three buckets: median of 10, 5, 0 dangerous permissions
Communications and business apps may require more
"dangerous" permissions for their core functionality
Medical apps have fewer total permissions but fewer alternatives

Correlation between total and dangerous permissions
Buckets of dangerous permissions → bundled?
Communications clearly visible top right (green)

Linear Regression

Model 1

Response: Dangerous Permissions

Predictors:
    Category

Significant:
    Category: $p \ll 0.05$

R^2: 0.24

Best: Racing, Sports Games, Brain
Worst: Communication, Medical,
        Social

Model 2 ⭐

Response: Dangerous Permissions

Predictors:
        Total Permissions
        Category

Significant:
    Total Permissions: $p \ll 0.05$
    Category: $p \ll 0.05$

R^2: 0.93

Best: Libraries, Productivity, Social
Worst: Finance, Lifestyle, Health
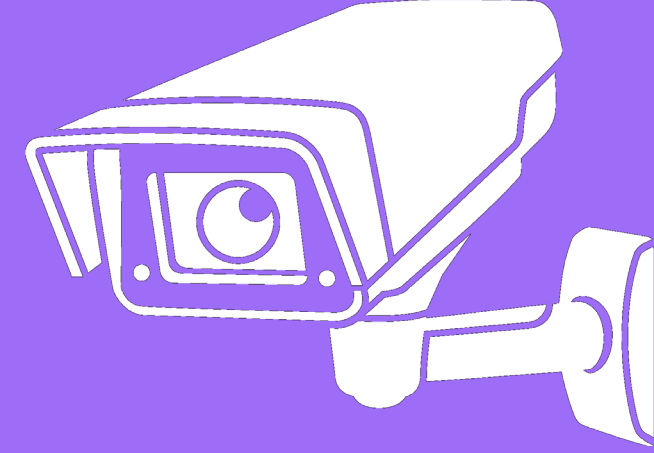
Model 3

Response: Dangerous Permissions

Predictors:
        Total Permissions
        Category
        Interaction

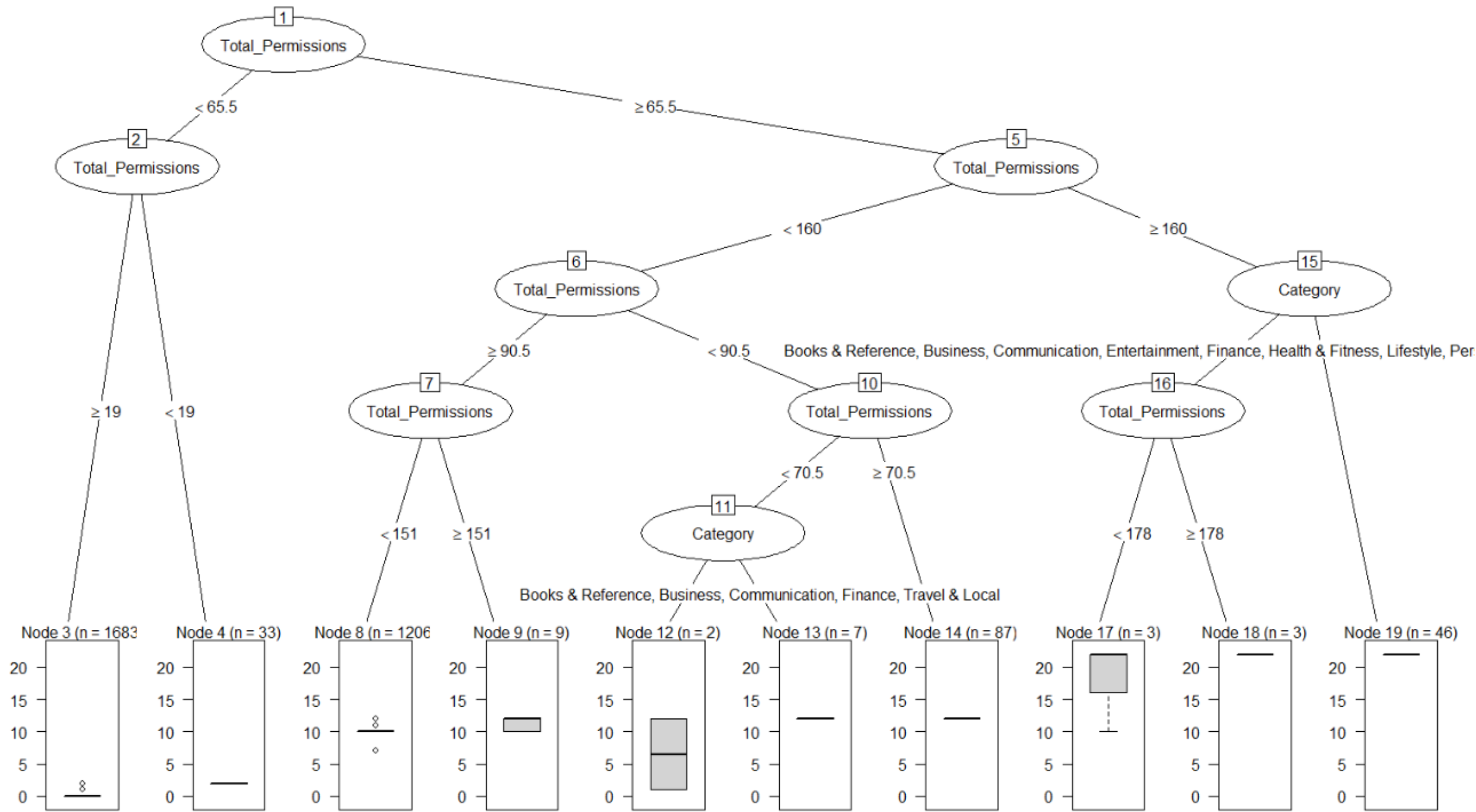Significant:
    All three: $p \ll 0.05$

R^2: 0.93

Best*: Weather, Travel,
Personaliz'n
Worst*: Racing, Comms, Medical

Project Overview   Data Summary   **Model and Analysis**   Recommendations
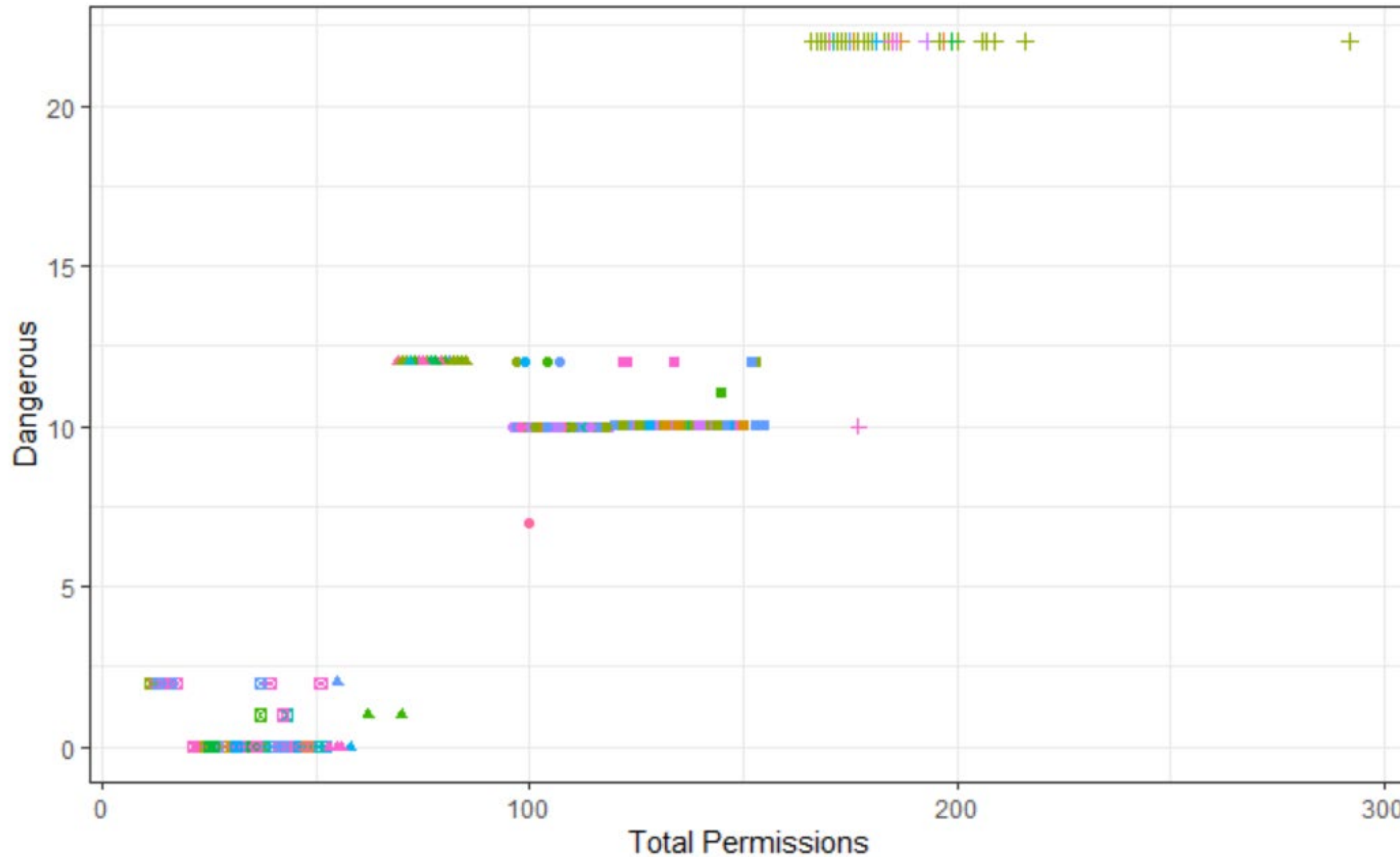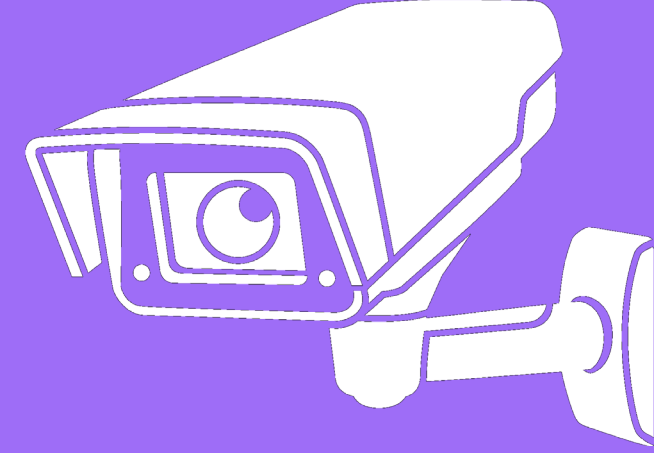
Decision Tree

Number of dangerous permissions is largely determined by the total number of permissions

Category really only matters when the total permissions are between 65 and 70

Clustering

Shape = Cluster, Color = Category

Three clear clusters (picked up with k = 3)

For fun, used k = 5

Clusters based on total permissions, not category

Category Analysis

Category	Total Permissions	Dangerous Permissions	Core Functionality	Competition
Personalization	+	+	+	+
Books & Reference	+	+	-	+
Entertainment	+	+/-	-	+
Tools	+	+/-	-	+
Lifestyle	+	-	-	+
Medical	+	-	+	-
Communication	-	-	+	-
Business	-	-	+	-
Weather	-	+	-	-

Best/Worst Apps

Package	Total_Permissions	Dangerous	Category	Freq	Score
com.lge.lglink.group	216	22	Entertainment	278	323
com.sks.android.joopasoo	199	22	Finance	54	308
com.nhn.android.search	197	22	Books & Reference	54	306
com.IQzone.PicDial	193	22	Social	132	302
com.econtacts.main	186	22	Social	132	295
mobilehome.backup	185	22	Tools	280	292
com.bibiku	184	22	Tools	280	291
com.skypenightmare	181	22	Social	132	290
com.wroclawstudio.screencaller	181	22	Personalization	142	290
com.wifivoiper	181	22	Travel & Local	144	290

Package	Total_Permissions	Dangerous	Category	Freq	Score
com.poidio.ServiceViewer	21	0	Tools	280	18.2
com.socialnmobile.hangulkeyboard	11	2	Tools	280	18.2
com.icogno.cleverbot	11	2	Entertainment	278	18.2
com.themythos.emoticonkeyboard	11	2	Communication	249	18.5
com.asigbe.slidekeyboardpro	12	2	Tools	280	19.2
com.curvefish.apps.processmanager	22	0	Tools	280	19.2
com.greatbytes.fastreboot	22	0	Tools	280	19.2
com.jamesots.android.closeeverything2	22	0	Tools	280	19.2
com.klye.inputmethod.latin	12	2	Tools	280	19.2
com.menny.android.anysoftkeyboard	12	2	Tools	280	19.2

Score = Total + 5*Dangerous - 0.01*Apps in Category

Suggestions to Manage Your Privacy
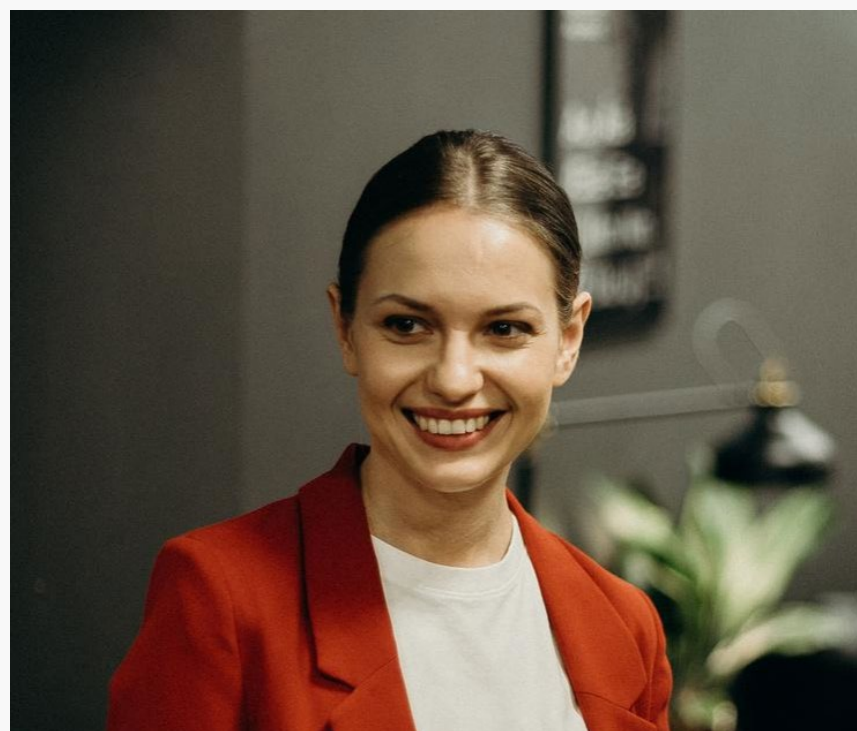
For apps in pretty safe categories:

You're probably okay!

For apps in more dangerous categories:

Consider doing research and choosing one with fewer permissions

If you can't find any detailed info on an app:

Remember total permission is the best predictor!

Questions?