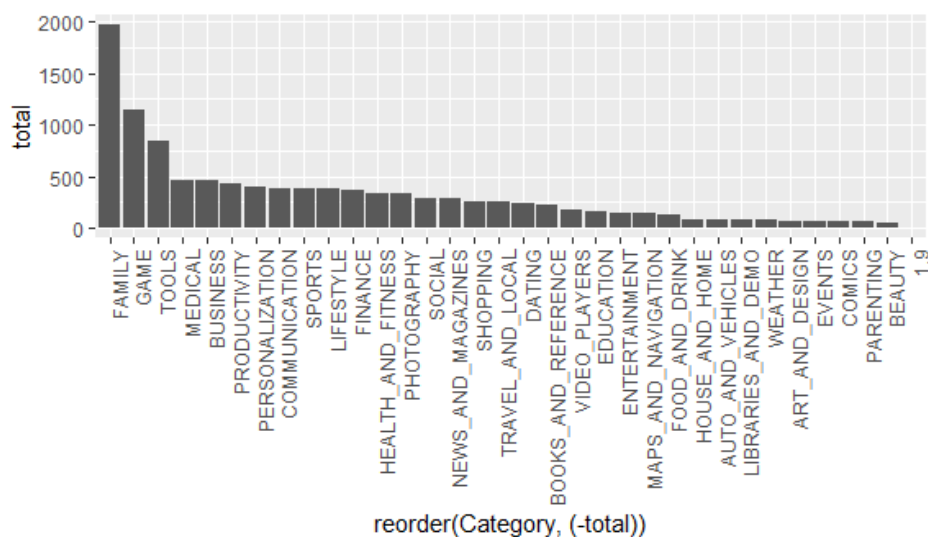


## LEB 3

# Distributions

DATASET : googleplaystore.csv

### Categorical variable



ภาพนี้แสดงให้เห็นถึงประเภทของเกมใน Google Playstore ซึ่งเกมใน Google Playstore มีทั้งหมด 34 Category เพื่อนำมาวิเคราะห์ข้อมูลเกี่ยวกับเกมประเภทไหนที่มีประเภทการดาวน์โหลดมากที่สุดได้ใน Google Playstore คืออะไร ซึ่งประเภทการดาวน์โหลดมากที่สุดจากในกราฟและข้อมูลคือ Family

Category	sum
1.9	1
ART_AND_DESIGN	65
AUTO_AND_VEHICLES	85
BEAUTY	53
BOOKS_AND_REFERENCE	231
BUSINESS	460
COMICS	60
COMMUNICATION	387
DATING	234
EDUCATION	156
ENTERTAINMENT	149
EVENTS	64
FAMILY	1972
FINANCE	366
FOOD_AND_DRINK	127
GAME	1144
HEALTH_AND_FITNESS	341
HOUSE_AND_HOME	88
LIBRARIES_AND_DEMO	85
LIFESTYLE	382
MAPS_AND_NAVIGATION	137
MEDICAL	463
NEWS_AND_MAGAZINES	283
PARENTING	60
PERSONALIZATION	392
PHOTOGRAPHY	335
PRODUCTIVITY	424
SHOPPING	260
SOCIAL	295
SPORTS	384
TOOLS	843
TRAVEL_AND_LOCAL	258
VIDEO_PLAYERS	175
WEATHER	82

Type of distributions : Frequency distribution – Bar chart

Values : จำนวน Category ที่มีประเภทการดาวน์โหลดมากที่สุดคือ Family

ยอดการดาวน์โหลดเท่ากับ 1972 ครั้ง

จำนวน Category ที่มีประเภทการดาวน์โหลดน้อยที่สุดคือ 1.9

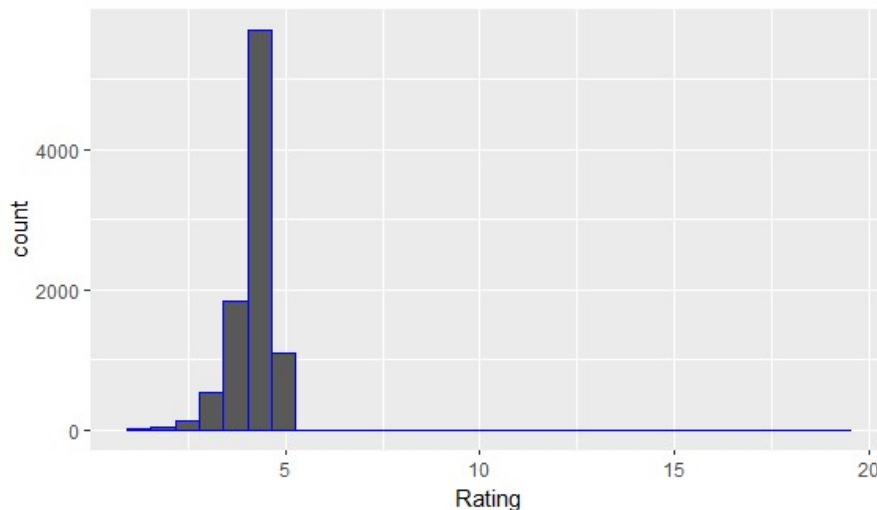
ยอดการดาวน์โหลดเท่ากับ 1 ครั้ง

Skewness : none เนื่องจากเรียงลำดับจากมากไปน้อยกราฟจึงดูเหมือนจะเบ้ขวา แต่จริงๆแล้วไม่ใช่ ข้อมูลกระจายตัวกัน

Outliers : none

Modalities : none

## Numerical variable



จากรูปแสดงให้เห็นถึง Overall user rating ของแต่ละแอปใน Google playstore

Type of distributions : Frequency distribution – Histogram

	Max	Min	Mean	Median	Q1	Q3	IQR
Values :	19	1	4.193338	4.3	4	4.5	0.5

Overall user rating ของแต่ละแอปใน Google playstore ที่น้อยที่สุด คือ 1

Overall user rating ของแต่ละแอปใน Google playstore ที่มากที่สุด คือ 19

ค่าเฉลี่ย(Mean) ประมาณ 4.19

ค่ามัธยฐาน(Median) คือ 4.3

จาก  $IQR = Q3 - Q1$  จะได้  $IQR = 0.5$

Skewness : Negatively skewed (เบ้ซ้าย) ดูจาก  $Median > Mean$  ที่ไม่ได้เอา

Mode มาคิดเพราะในข้อมูลจริง Mode มันไม่มีจริง!

Outliers : มี Outliers อย่างเห็นได้ชัด

Modalities : Unimodal