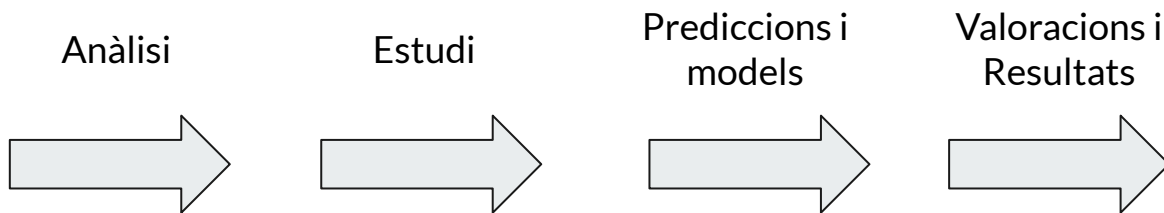# A quines aplicacions de GoogleStore he de posar anuncis?

Marc Bosom Saperas
NIU: 1606776
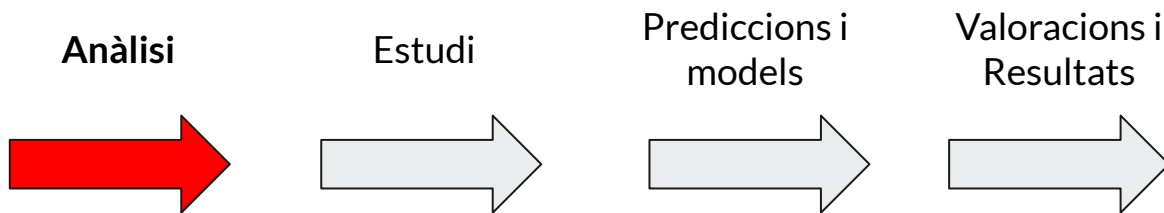
Plantejament

Anàlisi

Estudi

Prediccions i models

Valoracions i Resultats

Plantejament

Anàlisi   Estudi   Prediccions i models   Valoracions i Resultats

Anàlisi

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
0	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	4.1	159	19M	10,000+	Free	0	Everyone	Art & Design	January 7, 2018	1.0.0	4.0.3 and up
1	Coloring book moana	ART_AND_DESIGN	3.9	967	14M	500,000+	Free	0	Everyone	Art & Design;Pretend Play	January 15, 2018	2.0.0	4.0.3 and up
2	U Launcher Lite – FREE Live Cool Themes, Hide ...	ART_AND_DESIGN	4.7	87510	8.7M	5,000,000+	Free	0	Everyone	Art & Design	August 1, 2018	1.2.4	4.0.3 and up
3	Sketch - Draw & Paint	ART_AND_DESIGN	4.5	215644	25M	50,000,000+	Free	0	Teen	Art & Design	June 8, 2018	Varies with device	4.2 and up
4	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	4.3	967	2.8M	100,000+	Free	0	Everyone	Art & Design;Creativity	June 20, 2018	1.1	4.4 and up

# Anàlisi

- MIDA: 10841x13 → 9660x13

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10841 entries, 0 to 10840
Data columns (total 13 columns):
 #   Column          Non-Null Count   Dtype
---  ------          --------------   -----
 0   App             10841 non-null   object
 1   Category        10841 non-null   object
 2   Rating          9367 non-null    float64
 3   Reviews         10841 non-null   object
 4   Size            10841 non-null   object
 5   Installs        10841 non-null   object
 6   Type            10840 non-null   object
 7   Price           10841 non-null   object
 8   Content Rating  10840 non-null   object
 9   Genres          10841 non-null   object
 10  Last Updated    10841 non-null   object
 11  Current Ver     10833 non-null   object
 12  Android Ver     10838 non-null   object
dtypes: float64(1), object(12)
memory usage: 1.1+ MB
```

Plantejament

Anàlisi

Estudi

Prediccions i models

Valoracions i Resultats
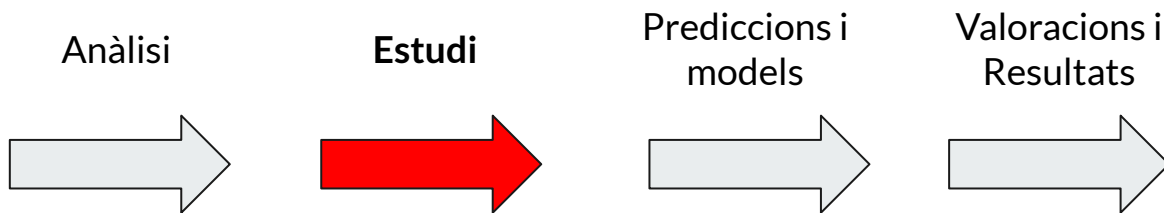
Estudi

1. Eliminació de NULLS
2. Modificació de les dades
3. Correlació de les dades

Plantejament de diferents preguntes:

1. Quines són les categories millor classificades?
2. Quines són les aplicacions més descarregades?
3. Quina és la franja d'edat amb més aplicacions?
4. Quines són les categories amb més aplicacions de pagament?

Modificació de les dades

```python
In [36]: dataset['Installs'] = dataset['Installs'].str.replace("+", "")
         dataset['Installs'] = dataset['Installs'].str.replace(",", "")
```

```python
In [37]: dataset['Price']= dataset["Price"].str.replace("$", "")
```

```python
In [38]: dataset['Size'] = dataset["Size"].str.replace("M", "")
         dataset['Size'] = dataset['Size'].str.replace("Varies with device","-1")
```
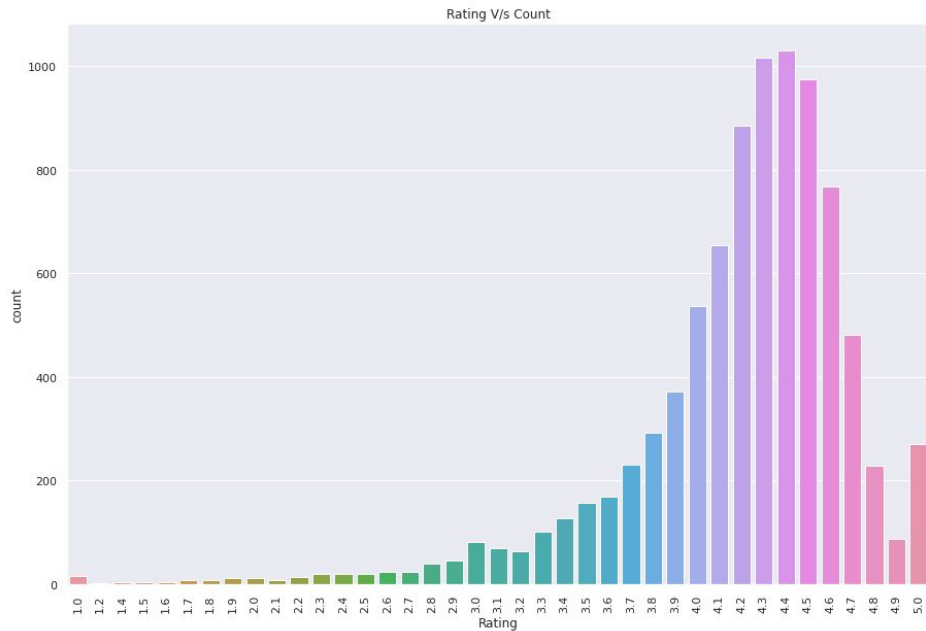
```python
In [39]: dataset['Size'] = dataset['Size'].apply(lambda x: str(round((float(x.rstrip('k'))/1024),1) )if x[-1]=='k' else x)
```
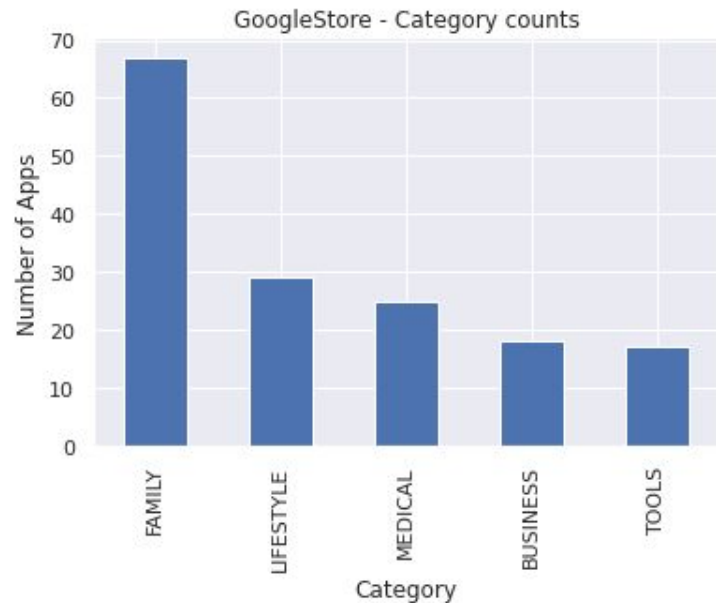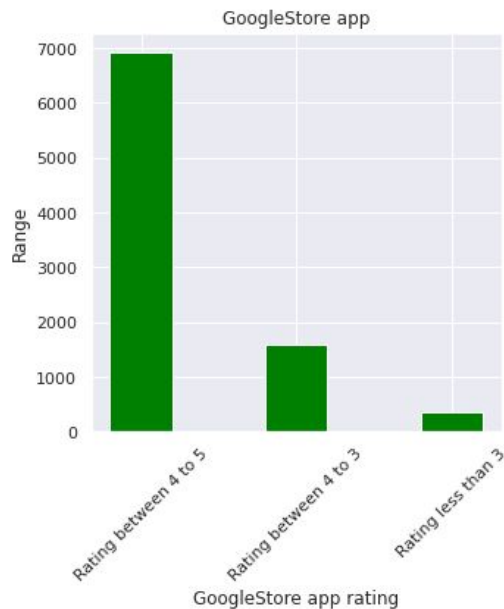
Correlació de les dades

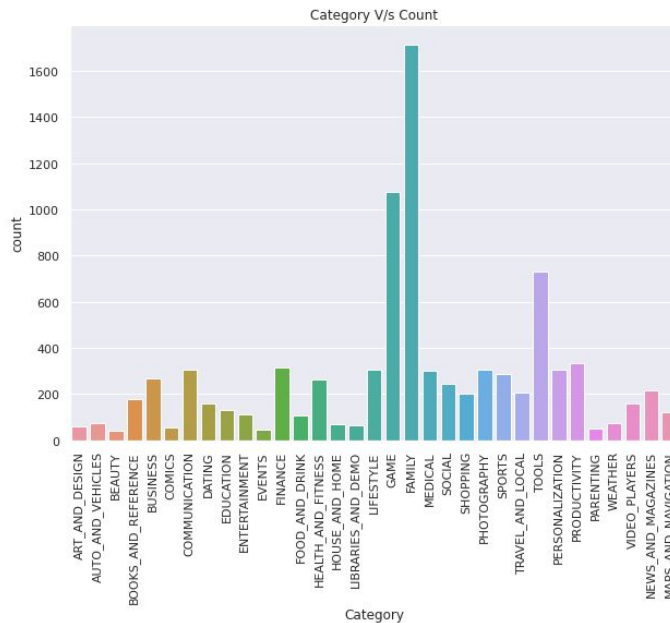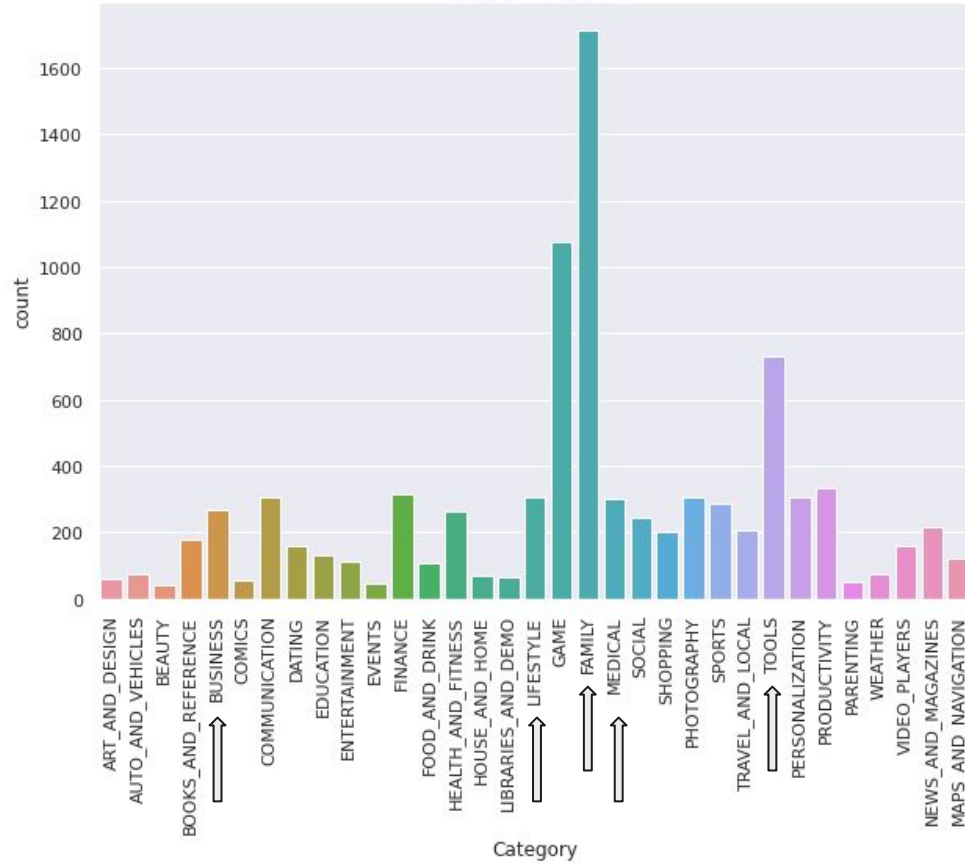# Quines són les categories millor classificades?

Rating V/s Count

Quines són les categories millor classificades?

Quines són les categories millor classificades?

Category V/s Count

Category V/s Count

Quines són les aplicacions més descarregades?

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
386	Hangouts	COMMUNICATION	4.0	3419433.0	-1.0	1.000000e+09	Free	0.0	Everyone	Communication	July 21, 2018	Varies with device	Varies with device
3736	Google News	NEWS_AND_MAGAZINES	3.9	877635.0	13.0	1.000000e+09	Free	0.0	Teen	News & Magazines	August 1, 2018	5.2.0	4.4 and up
4098	Maps - Navigate & Explore	TRAVEL_AND_LOCAL	4.3	9231613.0	-1.0	1.000000e+09	Free	0.0	Everyone	Travel & Local	July 31, 2018	Varies with device	Varies with device
3909	Instagram	SOCIAL	4.5	66509917.0	-1.0	1.000000e+09	Free	0.0	Teen	Social	July 31, 2018	Varies with device	Varies with device
2604	Instagram	SOCIAL	4.5	66577446.0	-1.0	1.000000e+09	Free	0.0	Teen	Social	July 31, 2018	Varies with device	Varies with device

```
In [55]: len(dataset[dataset["Installs"]>= 1000000000])
Out[55]: 49
```

Quines són les aplicacions més descarregades?

```
In [56]: top_10_installed_and_rated_apps = dataset.sort_values(by=["Installs", "Rating"], ascending=False)
         top_10_installed_and_rated_apps.head()
```

Out[56]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
1654	Subway Surfers	GAME	4.5	27722264.0	76.0	1.000000e+09	Free	0.0	Everyone 10+	Arcade	July 12, 2018	1.90.0	4.1 and up
1700	Subway Surfers	GAME	4.5	27723193.0	76.0	1.000000e+09	Free	0.0	Everyone 10+	Arcade	July 12, 2018	1.90.0	4.1 and up
1750	Subway Surfers	GAME	4.5	27724094.0	76.0	1.000000e+09	Free	0.0	Everyone 10+	Arcade	July 12, 2018	1.90.0	4.1 and up
1872	Subway Surfers	GAME	4.5	27725352.0	76.0	1.000000e+09	Free	0.0	Everyone 10+	Arcade	July 12, 2018	1.90.0	4.1 and up
2545	Instagram	SOCIAL	4.5	66577313.0	-1.0	1.000000e+09	Free	0.0	Teen	Social	July 31, 2018	Varies with device	Varies with device

# Quines són les aplicacions més descarregades?

```
In [57]:  top_10_installed_and_reviewed_apps = dataset.sort_values(by=["Installs", "Reviews"], ascending=False)
          top_10_installed_and_reviewed_apps.head()
```

Out[57]:

	App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Rating	Genres	Last Updated	Current Ver	Android Ver
2544	Facebook	SOCIAL	4.1	78158306.0	-1.0	1.000000e+09	Free	0.0	Teen	Social	August 3, 2018	Varies with device	Varies with device
3943	Facebook	SOCIAL	4.1	78128208.0	-1.0	1.000000e+09	Free	0.0	Teen	Social	August 3, 2018	Varies with device	Varies with device
336	WhatsApp Messenger	COMMUNICATION	4.4	69119316.0	-1.0	1.000000e+09	Free	0.0	Everyone	Communication	August 3, 2018	Varies with device	Varies with device
3904	WhatsApp Messenger	COMMUNICATION	4.4	69109672.0	-1.0	1.000000e+09	Free	0.0	Everyone	Communication	August 3, 2018	Varies with device	Varies with device
2604	Instagram	SOCIAL	4.5	66577446.0	-1.0	1.000000e+09	Free	0.0	Teen	Social	July 31, 2018	Varies with device	Varies with device

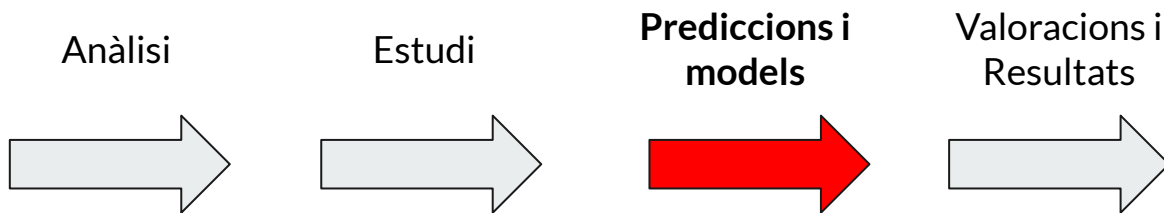# Quina és la franja d'edat amb més aplicacions?
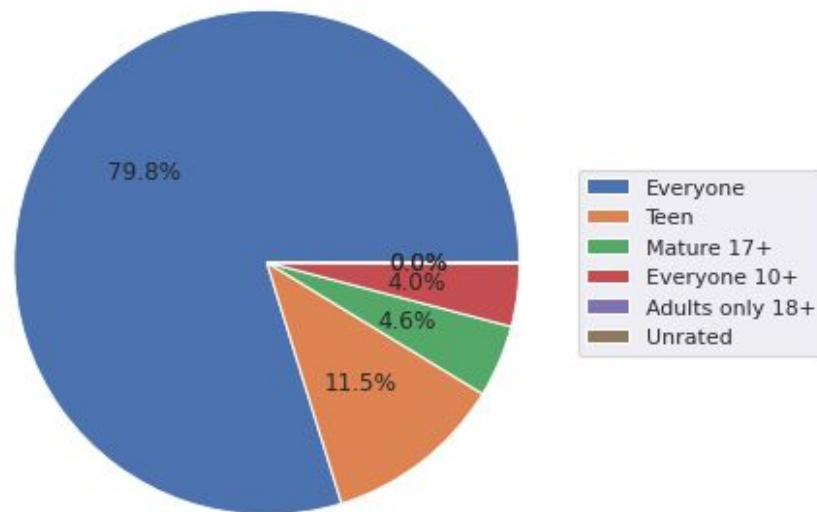
Quines són les categories amb més aplicacions de pagament?

Plantejament

Anàlisi

Estudi

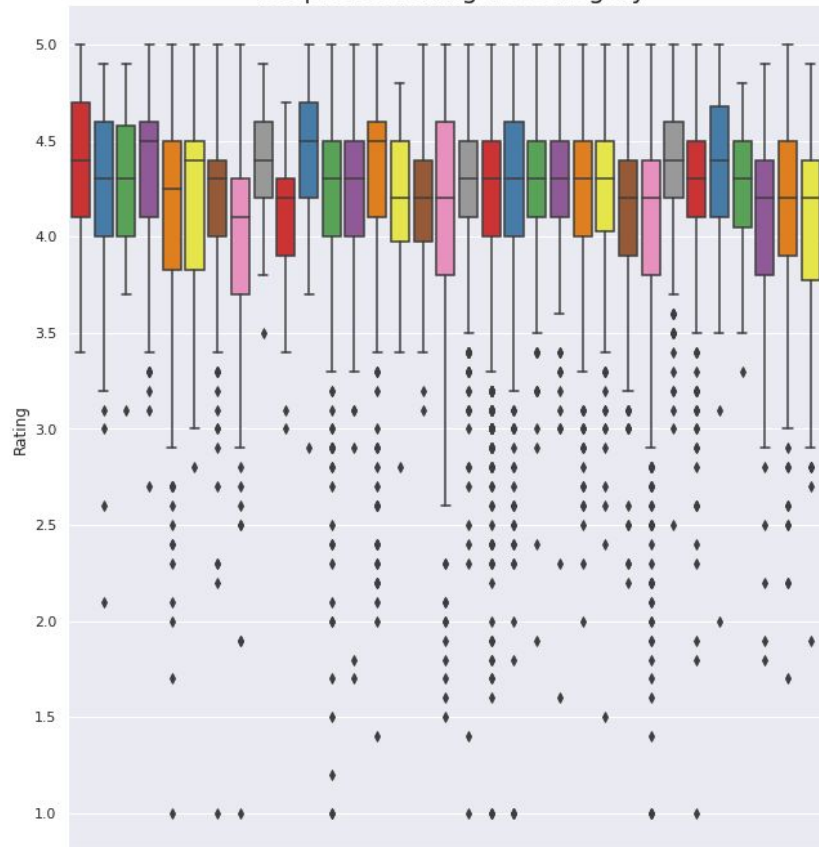**Prediccions i models**

Valoracions i Resultats
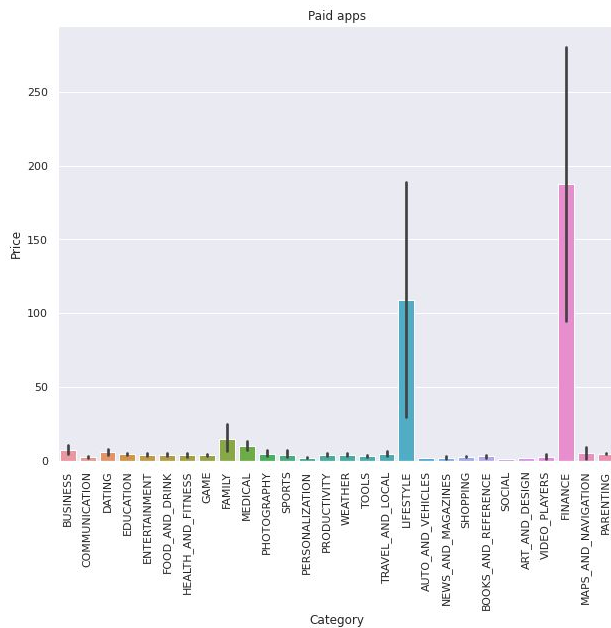
Prediccions

COM PUC ACONSEGUIR UNA ALTA VALORACIÓ?

CATEGORY

Boxplot of Rating VS Category

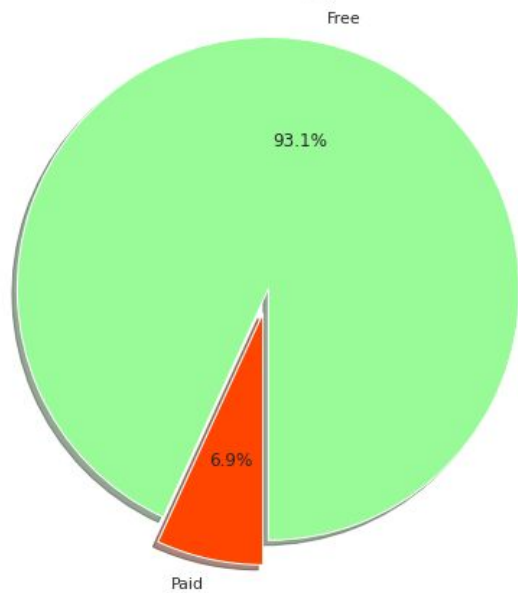# Reviews

Rating VS Reveiws

Size

Rating VS Size
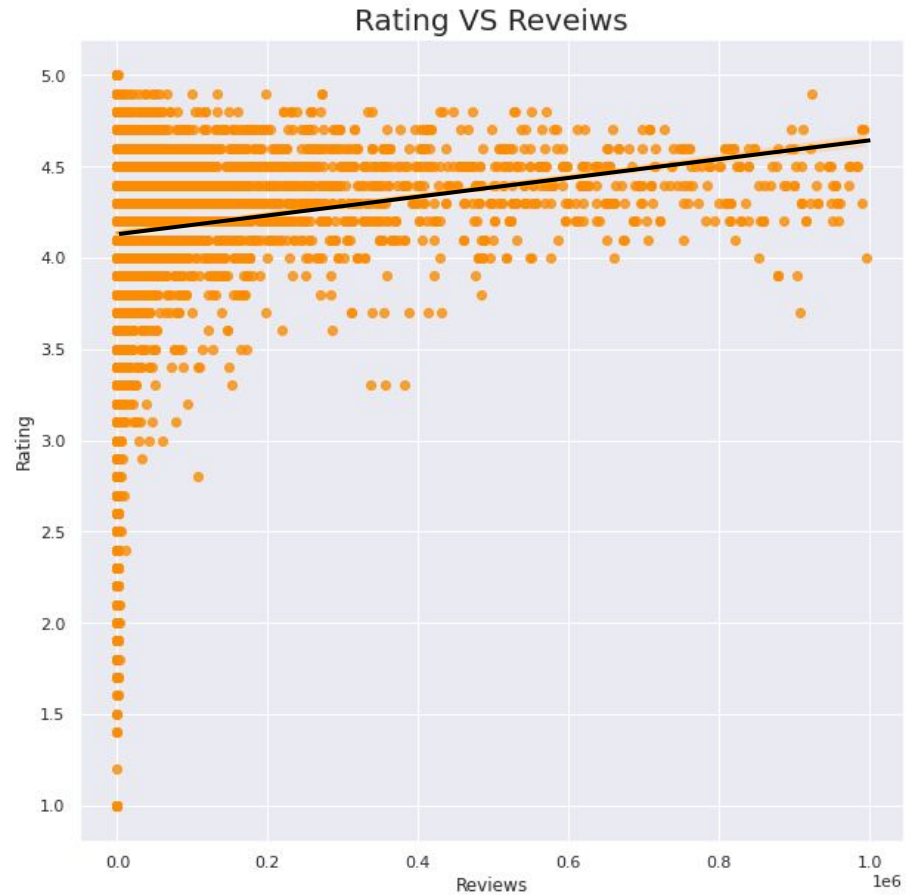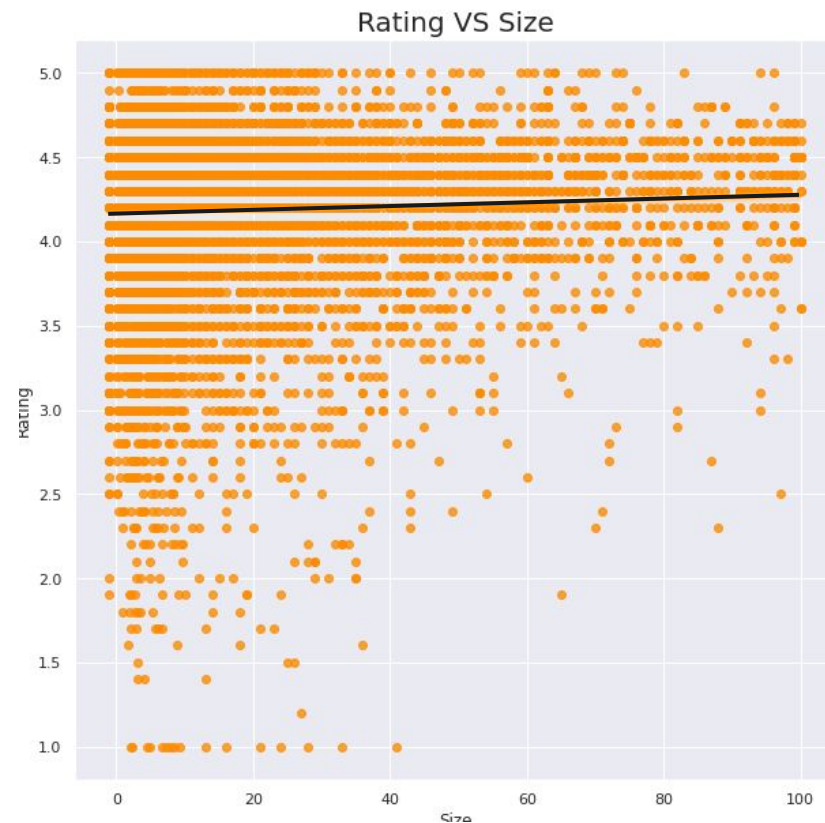
Installs

Rating VS Installs
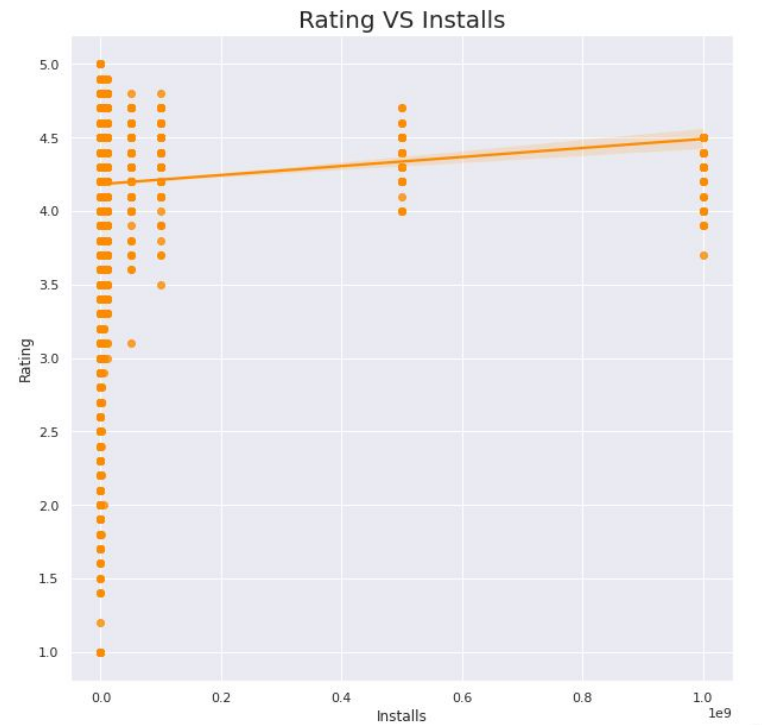
Type / Price

Percent of Free App in store


Free

93.1%

6.9%

Paid

Rating VS Price
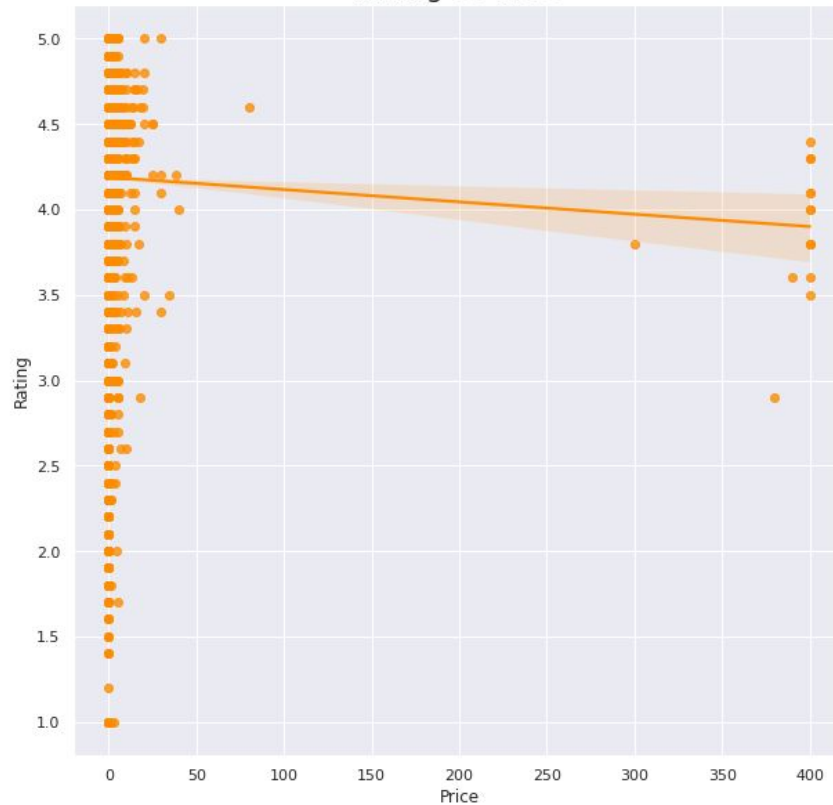
Type / Price

```
In [96]: dataset.loc[ dataset['Price'] == 0, 'PriceBand'] = '0 Free'
         dataset.loc[(dataset['Price'] > 0) & (dataset['Price'] <= 0.99), 'PriceBand'] = '1 cheap'
         dataset.loc[(dataset['Price'] > 0.99) & (dataset['Price'] <= 2.99), 'PriceBand']   = '2 not cheap'
         dataset.loc[(dataset['Price'] > 2.99) & (dataset['Price'] <= 4.99), 'PriceBand']   = '3 normal'
         dataset.loc[(dataset['Price'] > 4.99) & (dataset['Price'] <= 14.99), 'PriceBand']   = '4 expensive'
         dataset.loc[(dataset['Price'] > 14.99) & (dataset['Price'] <= 29.99), 'PriceBand']   = '5 pretty expensive'
         dataset.loc[(dataset['Price'] > 29.99), 'PriceBand']  = '6 too expensive'
```
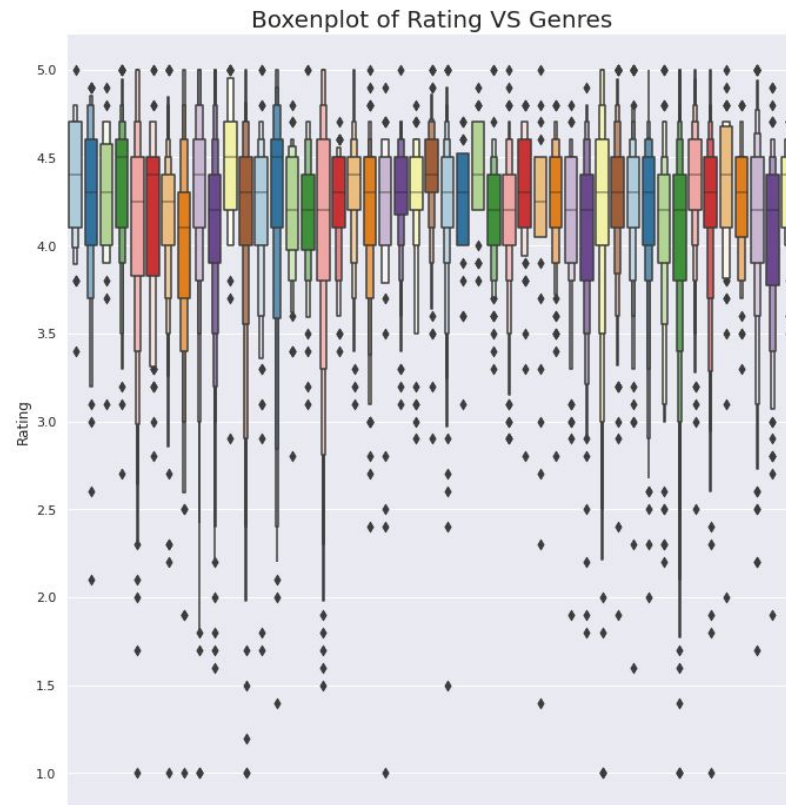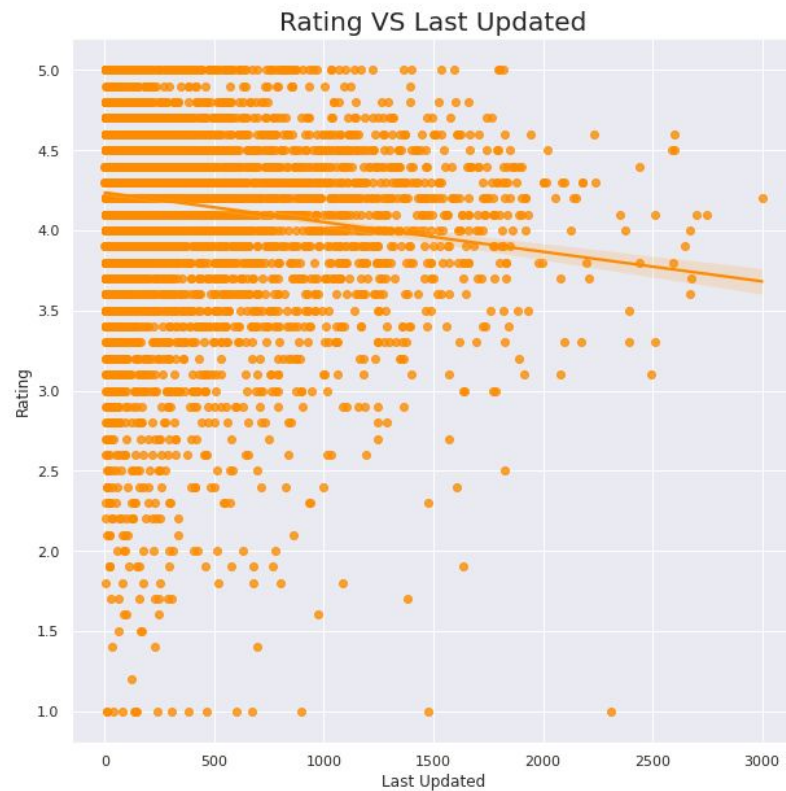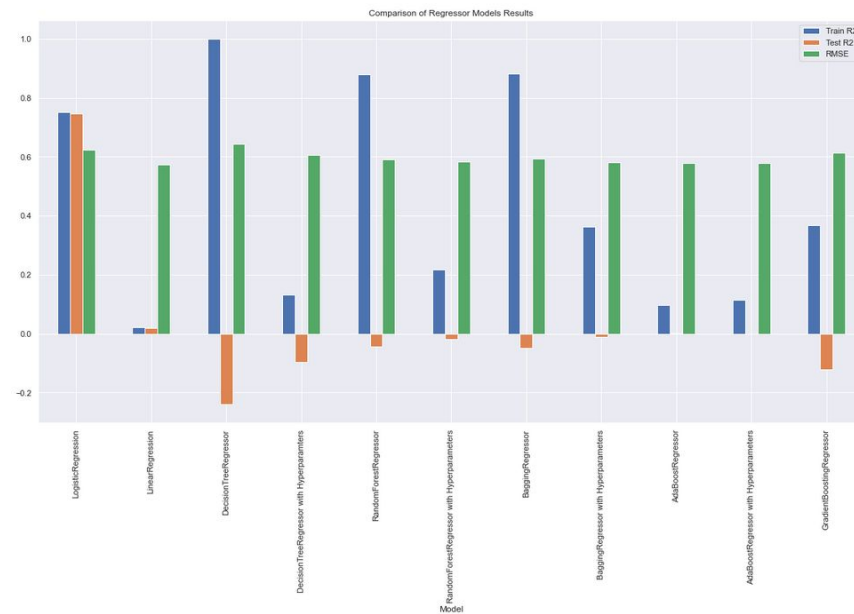
Type / Price

Boxen plot Rating VS PriceBand

Content Rating

Genres

Boxenplot of Rating VS Genres
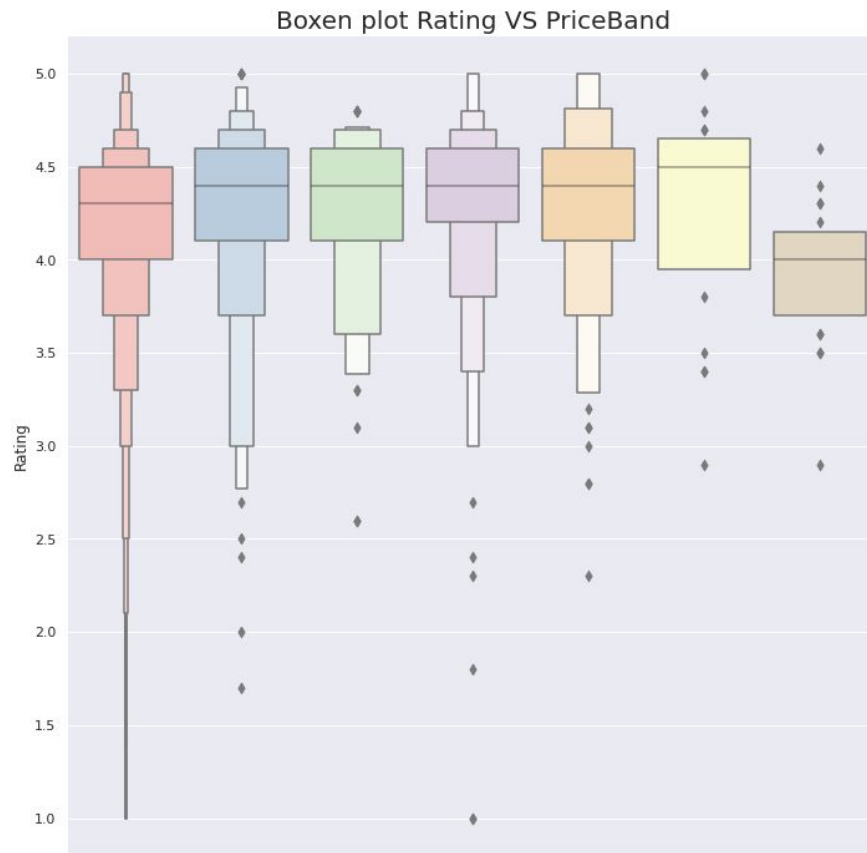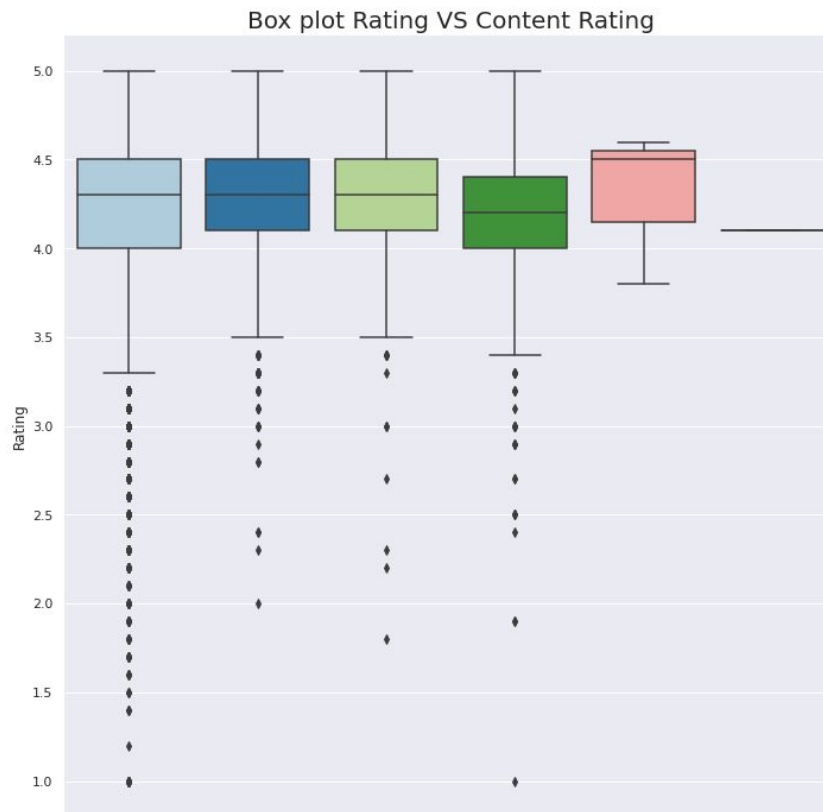
Last Updated

Rating VS Last Updated

Comparativa de Models

Plantejament

Anàlisi

Estudi

Prediccions i models

Valoracions i Resultats
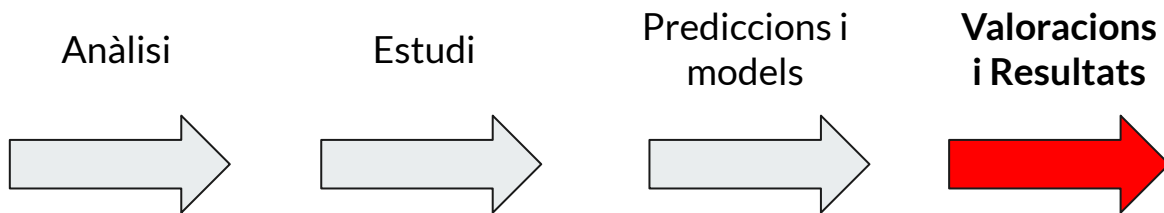
Valoracions i Resultats: on posem anuncis?

INSTALLS ⬅➡ REVIEWS

PRICE ⬇

RATING ➡ {
REVIEWS

SIZE

INSTALLS

LAST UPDATED
}

Logistic Regression ⬆

ALGUNA PREGUNTA?