# Importing Liabrearies

In [19]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [20]:

```
df = pd.read_csv("Top-Apps-in-Google-Play.csv")
```

In [21]:

```
df.head()
```

Out[21]:

| | Unnamed: 0 | App Name | App Id | Category | Developer Id | |
|---|---|---|---|---|---|---|
| **0** | 1 | Google Play services | com.google.android.gms | Tools | Google LLC | https://de |
| **1** | 2 | YouTube | com.google.android.youtube | Video Players & Editors | Google LLC | https://suppo |
| **2** | 3 | Google | com.google.android.googlequicksearchbox | Tools | Google LLC | h |
| **3** | 4 | Google Maps - Navigate & Explore | com.google.android.apps.maps | Travel & Local | Google LLC | |
| **4** | 5 | Google Text-to-Speech | com.google.android.tts | Tools | Google LLC | |

In [22]:

```
df.describe()
```

Out[22]:

| | Unnamed: 0 |
|---|---|
| **count** | 70.000000 |
| **mean** | 35.500000 |
| **std** | 20.351085 |
| **min** | 1.000000 |
| **25%** | 18.250000 |
| **50%** | 35.500000 |
| **75%** | 52.750000 |
| **max** | 70.000000 |

# Information about data

In [5]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 70 entries, 0 to 69
Data columns (total 10 columns):
 #   Column             Non-Null Count  Dtype 
---  ------             --------------  ----- 
 0   Unnamed: 0         70 non-null     int64 
 1   App Name           70 non-null     object
 2   App Id             70 non-null     object
 3   Category           70 non-null     object
 4   Developer Id       70 non-null     object
 5   Developer Website  70 non-null     object
 6   Developer Email    70 non-null     object
 7   Content Rating     70 non-null     object
 8   Ad Supported       70 non-null     bool  
 9   In App Purchases   70 non-null     bool  
dtypes: bool(2), int64(1), object(7)
memory usage: 4.6+ KB
```

# Checking for nan/missing values

In [23]:

```
df.isnull().sum()
```

Out[23]:

```
Unnamed: 0          0
App Name            0
App Id              0
Category            0
Developer Id        0
Developer Website   0
Developer Email     0
Content Rating      0
Ad Supported        0
In App Purchases    0
dtype: int64
```

# Top 5 No App Names

In [24]:

```
df["App Name"].head()
```

Out[24]:

```
0                Google Play services
1                             YouTube
2                              Google
3    Google Maps - Navigate & Explore
4               Google Text-to-Speech
Name: App Name, dtype: object
```

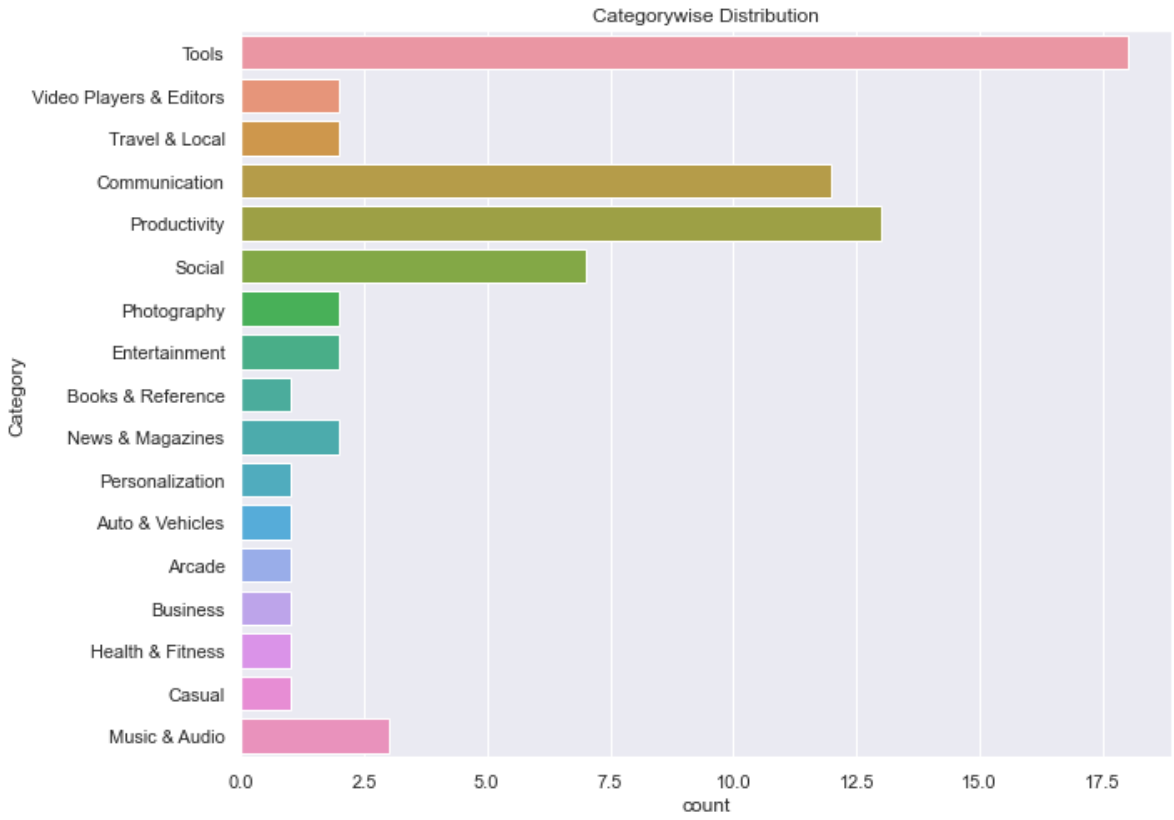
# Types of categories of apps exists on playstore

In [25]:

```
plt.figure(figsize=(10,8))
sns.countplot(data=df,y="Category")
plt.title("Categorywise Distribution")
```

Out[25]:

```
Text(0.5, 1.0, 'Categorywise Distribution')
```




# Types of developer id's on playstore

In [10]:

```
df["Developer Id"].value_counts()
```

Out[10]:

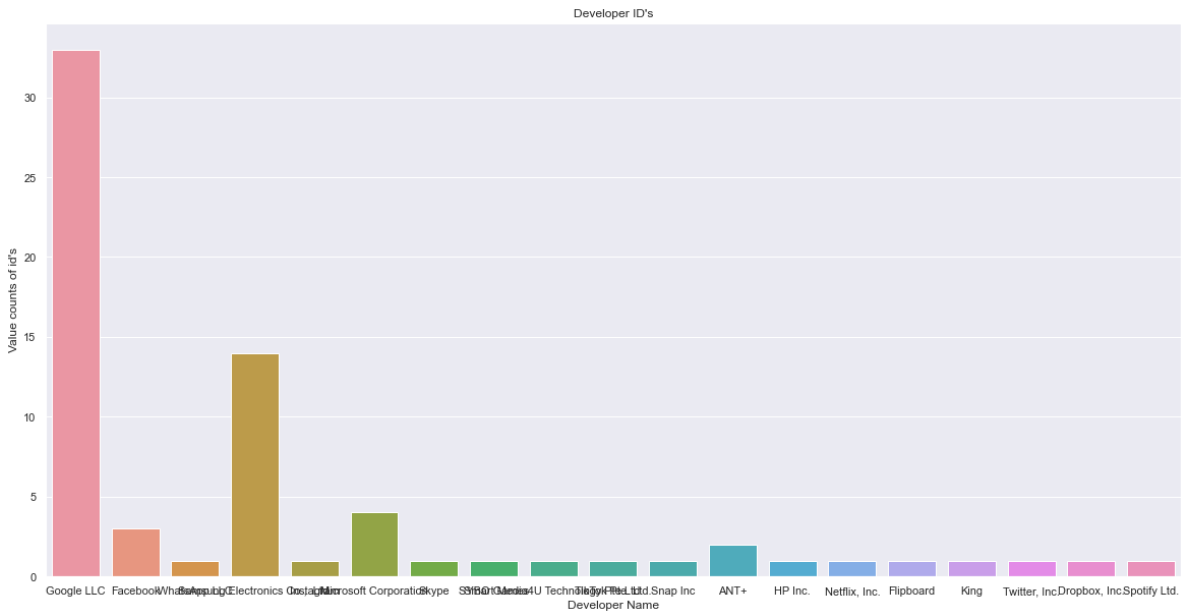
```
Google LLC                            33
Samsung Electronics Co.,  Ltd.        14
Microsoft Corporation                  4
Facebook                               3
ANT+                                   2
Dropbox, Inc.                          1
Twitter, Inc.                          1
King                                   1
Flipboard                              1
Netflix, Inc.                          1
HP Inc.                                1
TikTok Pte. Ltd.                       1
Snap Inc                               1
Smart Media4U Technology Pte.Ltd.      1
SYBO Games                             1
Skype                                  1
Instagram                              1
WhatsApp LLC                           1
Spotify Ltd.                           1
Name: Developer Id, dtype: int64
```

In [11]:

```
sns.set(style="darkgrid")
plt.figure(figsize=(20,10))
sns.countplot(x="Developer Id",data=df)
plt.title("Developer ID's")
plt.ylabel("Value counts of id's")
plt.xlabel("Developer Name")
```

Out[11]:

```
Text(0.5, 0, 'Developer Name')
```



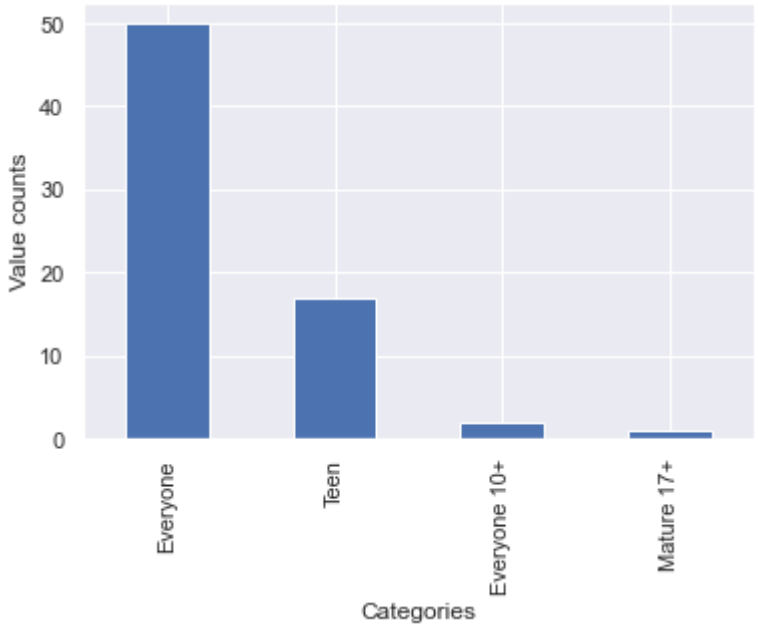

# Categoriwise content rating

In [12]:

```
df["Content Rating"].value_counts().plot(kind="bar")
plt.grid(True)
plt.xlabel("Categories")
plt.ylabel("Value counts")
```

Out[12]:

```
Text(0, 0.5, 'Value counts')
```



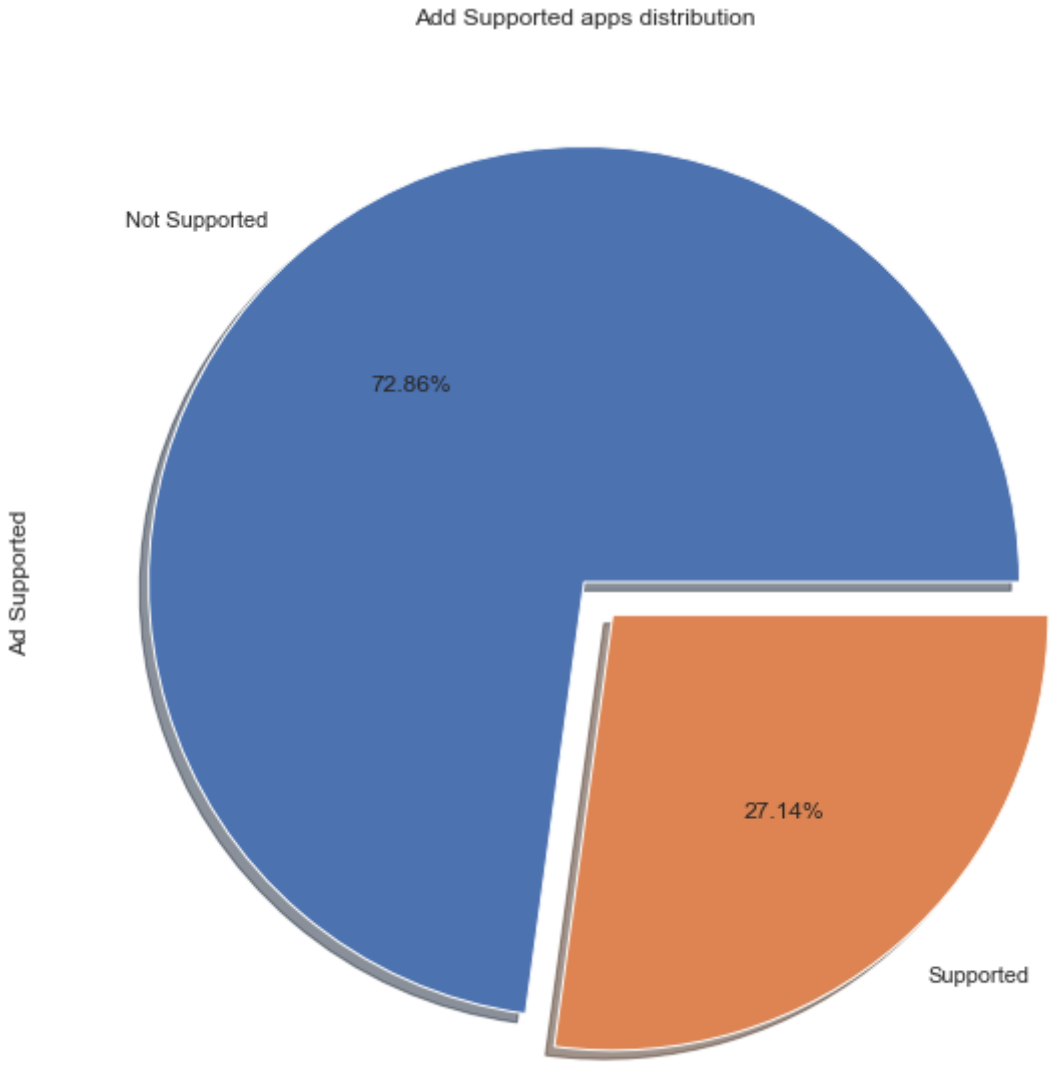

# Checking for add supported apps or not

In [13]:

```
plt.figure(figsize=(10,10))
x = ["Not Supported","Supported"]
df["Ad Supported"].value_counts().plot(kind="pie",autopct="%.2f%%",shadow=True,explode=(0,0
plt.title("Add Supported apps distribution")
```

Out[13]:

```
Text(0.5, 1.0, 'Add Supported apps distribution')
```



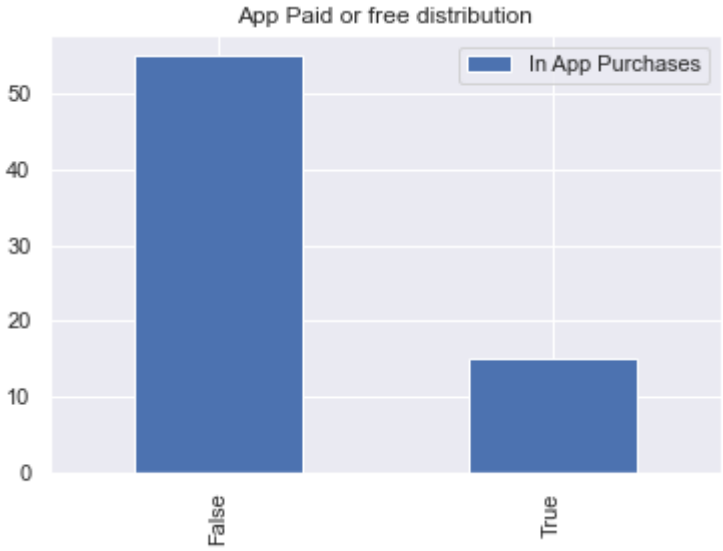

# Distribution of paid and free apps on playstore

In [18]:

```
df["In App Purchases"].value_counts().plot(kind="bar")
plt.legend()
plt.title("App Paid or free distribution")
```

Out[18]:

```
Text(0.5, 1.0, 'App Paid or free distribution')
```



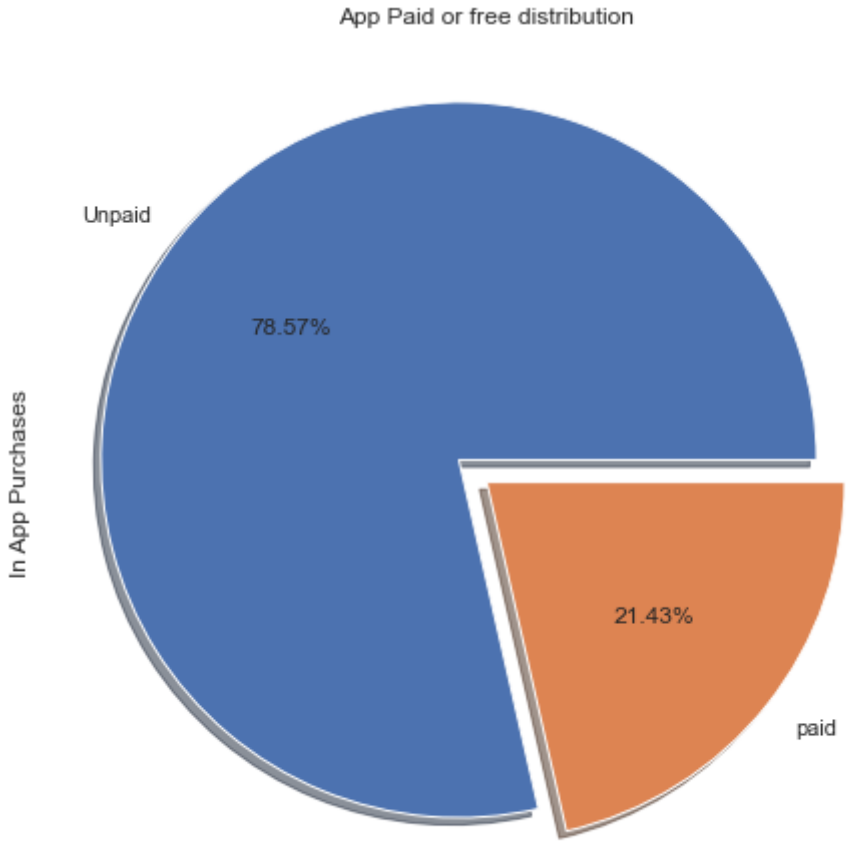
In [32]:

```
plt.figure(figsize=(8,10))
names = ["Unpaid","paid"]
df["In App Purchases"].value_counts().plot(kind="pie",labels=names,autopct="%.2f%%",shadow=


plt.title("App Paid or free distribution")
```

Out[32]:

```
Text(0.5, 1.0, 'App Paid or free distribution')
```




# Conclusion

- This data represents diffrent types of apps categories exists on google play store.There are 70 rows and 10 columns are there.

---

- In this dataset most of the data has object data-type and remaining has numeric and boolean data-type.

---

- In this datasets no missing values presents.

---

- Top 5 Names of apps on play store are:
- 1 Google Play services

- 2 YouTube
- 3 Google
- 4 Google Maps - Navigate & Explore
- 5 Google Text-to-Speech

---

- There are 17 types of apps exists on playstore out of all most apps exists on playstore has Tools Category.

---

- Top 5 Developer Id's on Play-store

---

- Google LLC 33
- Samsung Electronics Co., Ltd. 14
- Microsoft Corporation 4
- Facebook 3
- ANT+ 2

---

- Only 27% developers run ads in his apps and remaining 73% apps doesn't.

---

- Most rating for apps given by all categories of peoples.

---

- 78% Apps are free to use to user on play store remaining 22% are paid apps.

In [ ]: