1.What is the most expensive app on the Play Store?

```
max_price = df_data['Price'].max()
max_price = df_data[df_data['Price'] == max_price]
max_price['App']
```

25] ✓ 0.0s Python

```
4367    I'm Rich - Trump Edition
Name: App, dtype: object
```

2.Which genre has the highest number of apps?

```
top_genre = df_data['Genres'].value_counts().idxmax()
print(top_genre)
```

✓ 0.0s Python

```
Tools
```

3.What is the average size of free vs. paid apps?

```
avg_size_free_vs_paid = df_data.groupby('Type')['Size'].mean()
print(avg_size_free_vs_paid)
```

✓ 0.0s Python

```
Type
Free    15.433058
Paid    14.190778
Name: Size, dtype: float64
```

4.What are the top 5 most expensive apps with a perfect rating (5)?

```
top = df_data[(df_data['Rating'] == 5)].nlargest(5, 'Price')[['App']]
top
```

✓ 0.0s Python

	App
7477	USMLE Step 2 CK Flashcards
5482	meStudying: AP English Lit
7204	TI-84 CE Graphing Calculator Manual TI 84
5237	Hey AJ! It's Saturday!
8287	AC DC Power Monitor

5.How many apps have received more than 50K reviews?

```
reviews = df_data[(df_data['Reviews'] >= 50000)].count()
print(reviews['Reviews'])
```

✓ 0.0s Python

```
1167
```

6.What is the average price of apps, grouped by genre and number of installs?

```
avg = df_data.groupby(['Genres', 'Installs'])['Price'].mean()
print(avg)
```

✓ 0.0s Python

```
Genres  Installs
Action  10          0.000
        50          0.995
        100         0.330
        500         0.000
        1000        0.000
                    ...
Word    100000      0.000
        1000000     0.000
        5000000     0.000
        10000000    0.000
        50000000    0.000
Name: Price, Length: 808, dtype: float64
```

7.How many apps have a rating higher than 4.7, and what is their average price?

```
more_than_47 = df_data[(df_data['Rating']>4.7)].count()
avg_price = df_data[(df_data['Rating']>4.7)]['Price'].mean()
print(more_than_47.iloc[0])
print(avg_price)
```

✓ 0.0s Python

```
510
0.3770392156862745
```

8.What is Google's estimated revenue from apps with 5,000,000+ installs?

```
more_than = df_data[(df_data['Installs']>5000000)]
revenue = (more_than['Installs'] * more_than['Price']).sum()
revenue = revenue * 0.3
print(revenue)
```

✓ 0.0s Python

```
2970000.0
```

9.What are the maximum and minimum sizes of free vs. paid apps?

```python
free_max = df_data[(df_data['Type'] == 'Free')]['Size'].max()
free_min = df_data[(df_data['Type'] == 'Free')]['Size'].min()
paid_max = df_data[(df_data['Type'] == 'Paid')]['Size'].max()
paid_min = df_data[(df_data['Type'] == 'Paid')]['Size'].min()
print('free max {} , free min {}'.format(free_max,free_min))
print('paid max {} , paid min {}'.format(paid_max,paid_min))
```

✓ 0.0s Python

```
free max 53.0 , free min 0.00830078125
paid max 53.0 , paid min 0.0166015625
```

10.Is there a correlation between an app's rating, number of reviews, size, and its price?

```python
cor = df_data[['Rating', 'Reviews', 'Size', 'Price']].corr()
cor
```

✓ 0.0s Python

	Rating	Reviews	Size	Price
Rating	1.000000	0.060164	-0.019350	-0.041592
Reviews	0.060164	1.000000	0.161564	-0.010594
Size	-0.019350	0.161564	1.000000	-0.017944
Price	-0.041592	-0.010594	-0.017944	1.000000

11.How many apps exist for each type (free/paid) across different content ratings?

```python
app_count = df_data.groupby(['Type', 'Content Rating']).count()
app_count['App']
```

✓ 0.0s Python

```
Type  Content Rating
Free  Adults only 18+       2
      Everyone           5506
      Everyone 10+        182
      Mature 17+          270
      Teen                670
      Unrated               2
Paid  Everyone            501
      Everyone 10+         24
      Mature 17+           15
      Teen                 39
Name: App, dtype: int64
```

12.How many apps are compatible with Android version 4.x?

```python
ver_4 = df_data[df_data['Android Ver'].str.startswith("4")].count()
print(ver_4.iloc[0])
```

✓ 0.0s Python

```
5232
```