# Video Game Developer Analysis

**Joseph Davis**

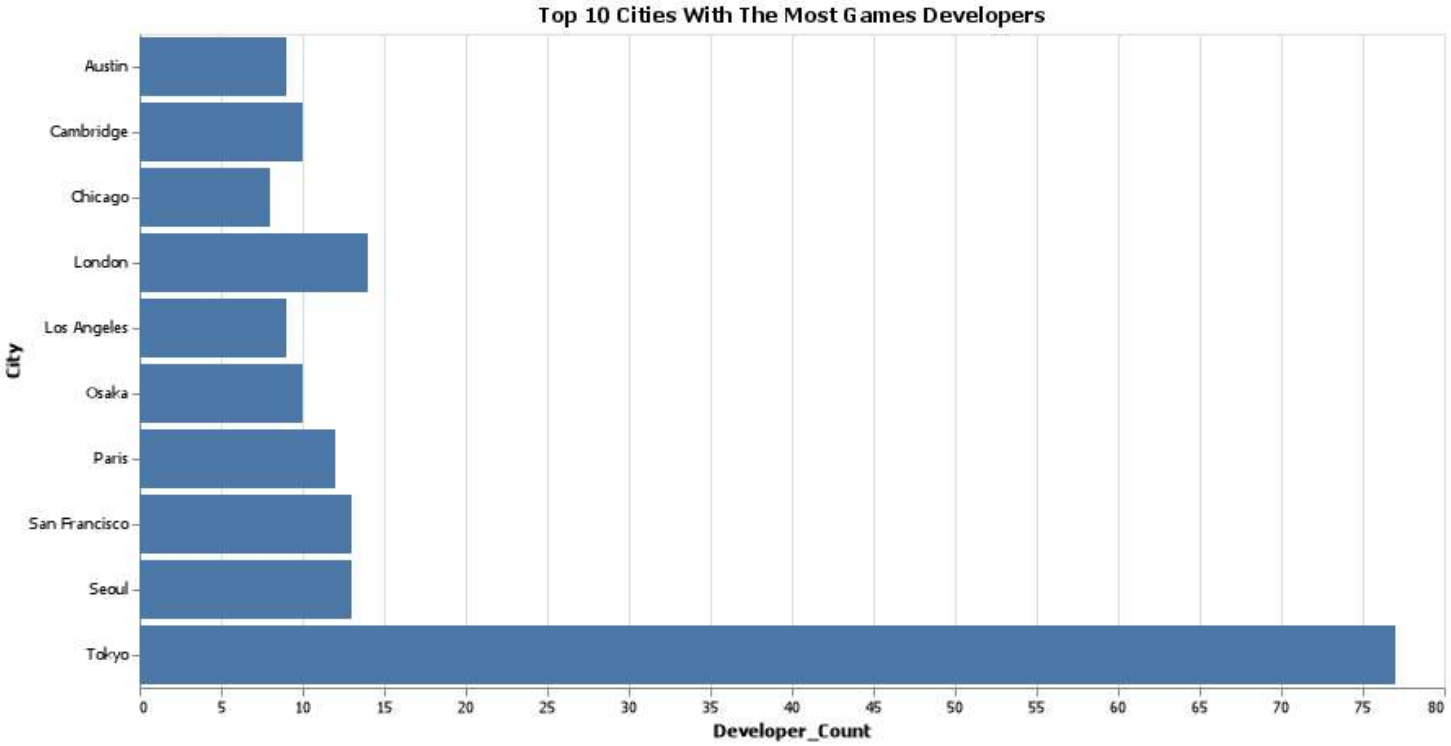
## Project Summary

This Project will focus on analyzing and visualizing different aspects regadring video game developers. This project will demonstrate skills using SQL quieries and data visualization using the Altair library in Python.

### Find the top 10 Cities with the most amount of game developers

```
SELECT DISTINCT city AS City, country AS Country, COUNT(developer) as Developer_Count
FROM developper
GROUP BY city
ORDER BY COUNT(developer) DESC
LIMIT 10
```

| City | Country | Developer_Count |
|---|---|---|
| Tokyo | Japan | 77 |
| London | England | 14 |
| San Francisco | US | 13 |
| Seoul | South Korea | 13 |
| Paris | France | 12 |
| Cambridge | US | 10 |
| Osaka | Japan | 10 |
| Los Angeles | US | 9 |
| Austin | US | 9 |
| Chicago | US | 8 |

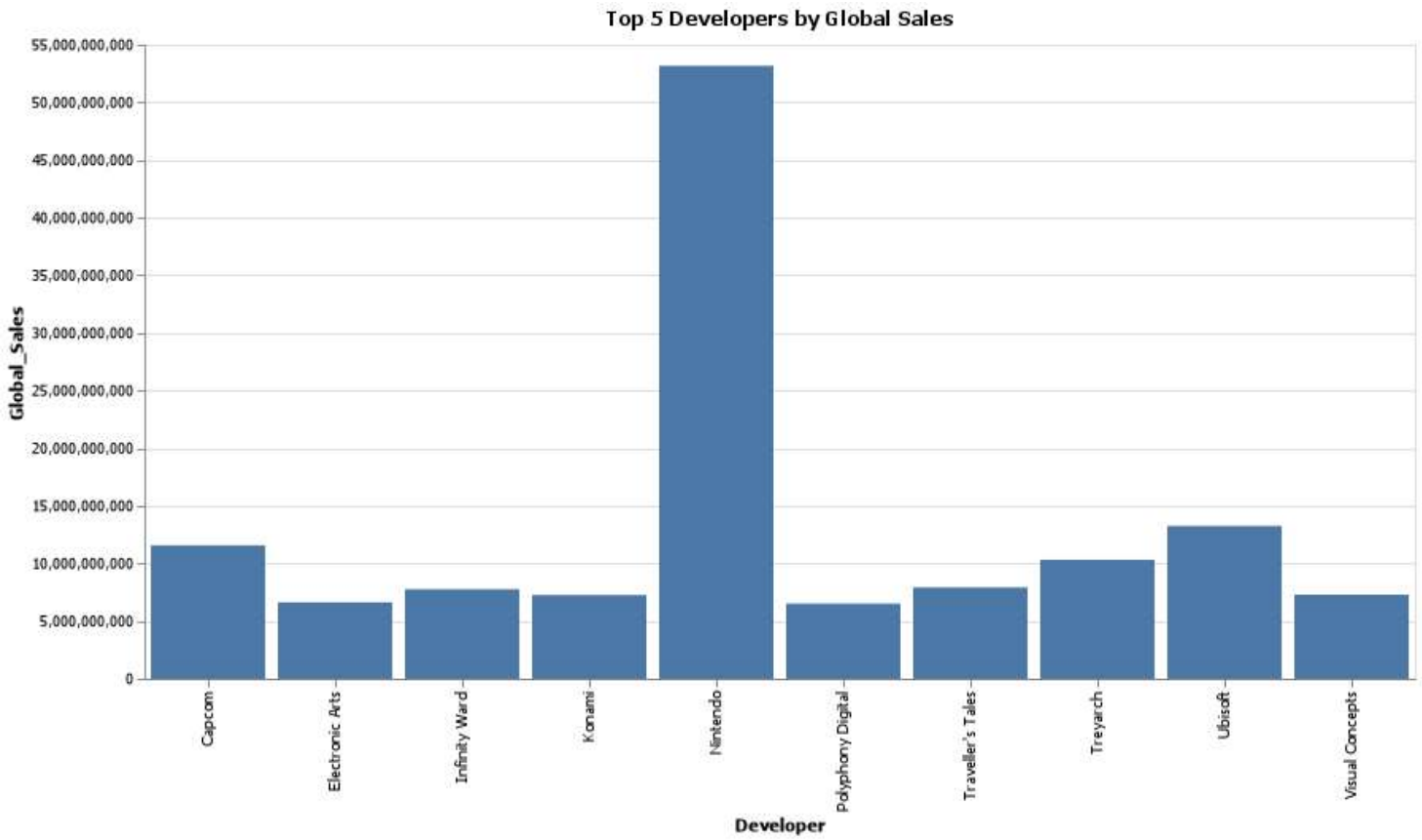



## Find The Top 10 Developers With The Highest mount Of Global Sales

```
SELECT d.developer AS Developer, sum(v.global_sales)*100000 AS Global_Sales
FROM vgsales v
JOIN developper d
ON  v.developer = d.developer
GROUP BY d.developer
ORDER BY Global_Sales DESC
LIMIT 10
```

| Developer | Global_Sales |
|---|---|
| Nintendo | 531710000 |
| Ubisoft | 132540000 |
| Capcom | 115710000 |
| Treyarch | 103160000 |
| Traveller's Tales | 79220000 |
| Infinity Ward | 77620000 |
| Visual Concepts | 72870000 |
| Konami | 72560000 |
| Electronic Arts | 66260000 |

| Developer | Global_Sales |
|---|---|
| Polyphony Digital | 65200000 |

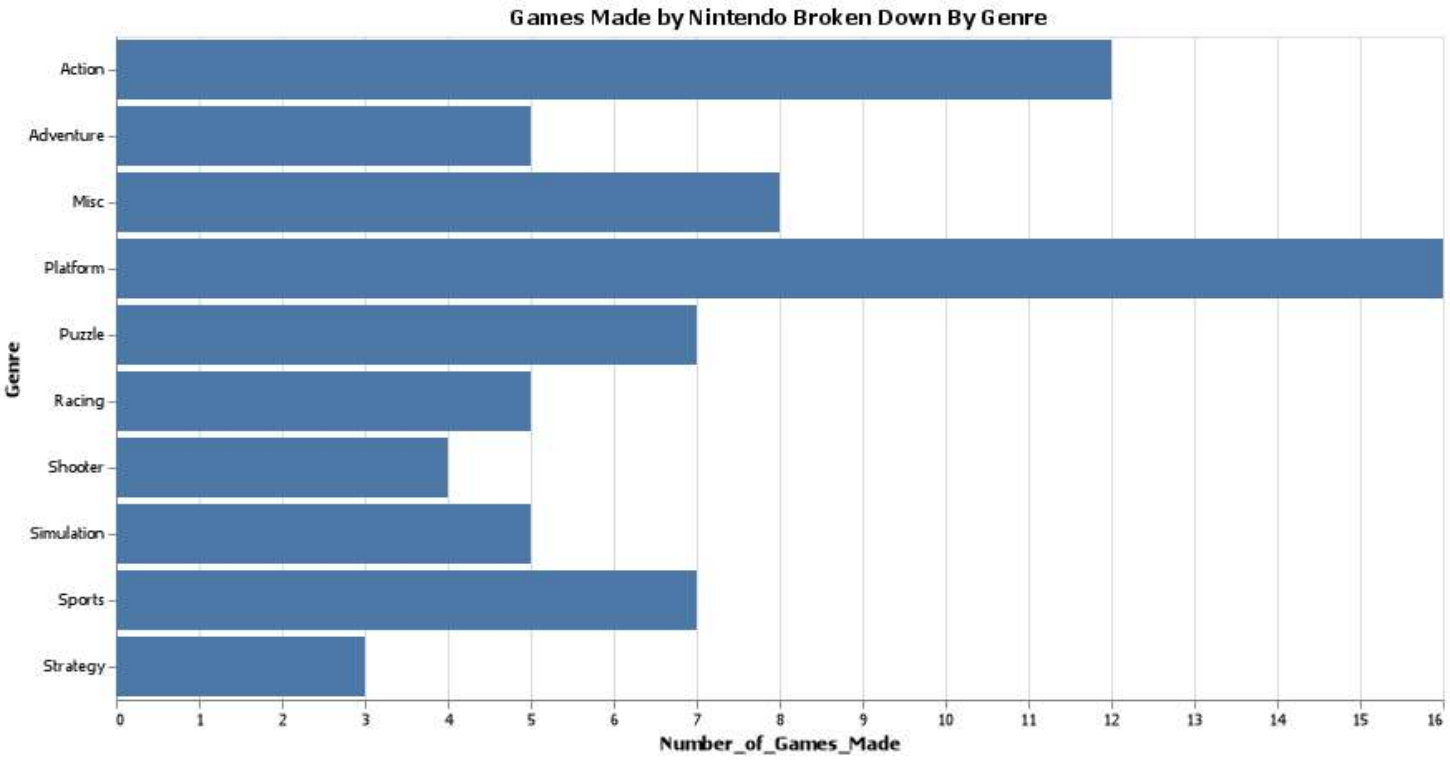



## Breakdown The Amount of Games Nintendo Has Developed By Genre

```
SELECT genre AS Genre, COUNT(GENRE) AS Number_of_Games_Made
FROM vgsales
WHERE developer = "Nintendo"
GROUP BY genre
ORDER BY COUNT(GENRE) ASC
```

| Genre | Number_of_Games_Made |
|---|---|
| Strategy | 3 |
| Shooter | 4 |
| Racing | 5 |
| Adventure | 5 |
| Simulation | 5 |
| Sports | 7 |

| Genre | Number_of_Games_Made |
|---|---|
| Puzzle | 7 |
| Misc | 8 |
| Action | 12 |
| Platform | 16 |

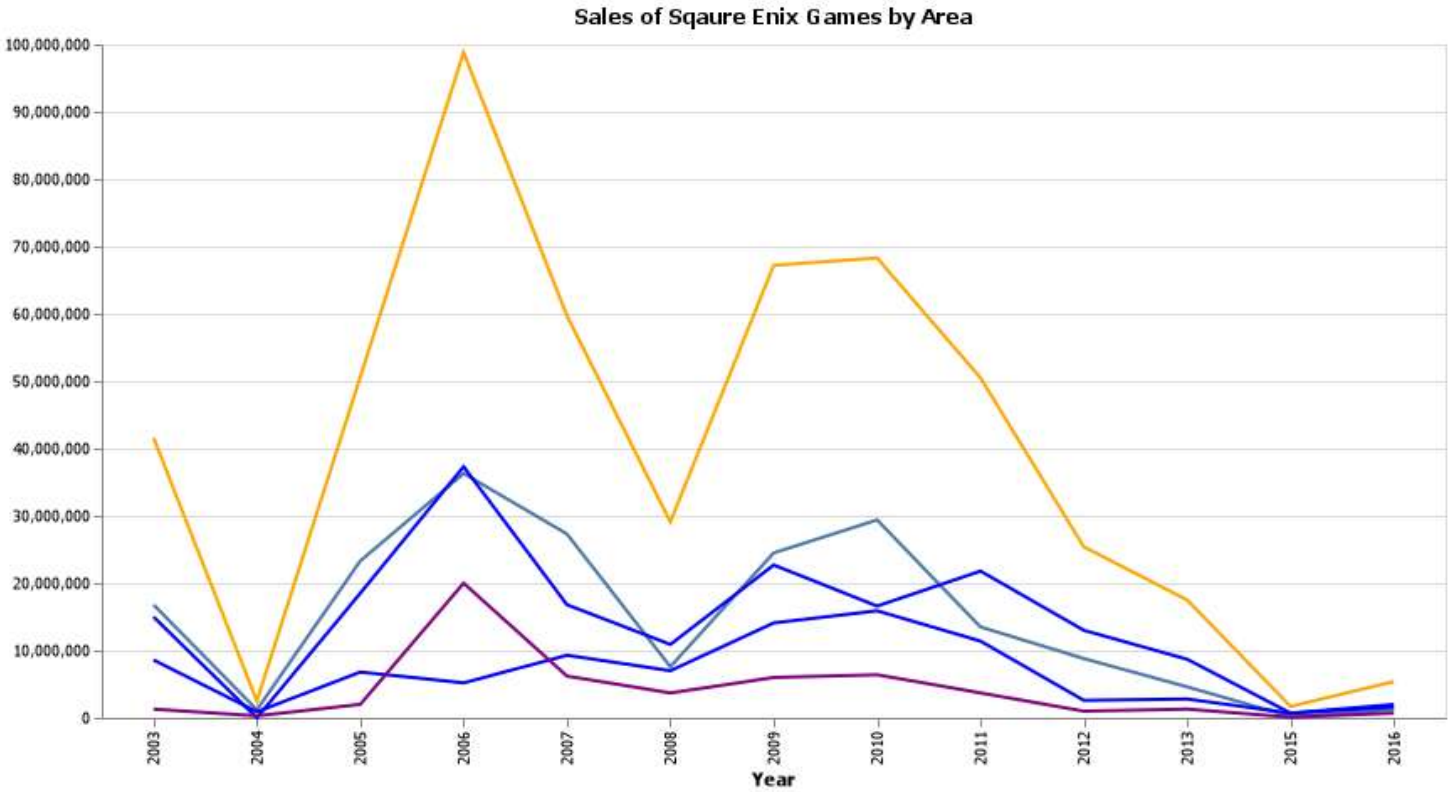



## Breakdown The Amount of Sales By Year of Square Enix (Devleoper), Include All Of The Sales By Area

```
SELECT year_of_release as Year,
    SUM(na_sales) AS North_America_Sales,
    SUM(eu_sales) AS Europe_Sales ,
    SUM(jp_sales) AS Japan_Sales,
    SUM(other_sales) AS Other_Sales,
    SUM(global_sales) AS Global_Sales
FROM vgsales
WHERE developer = "Square Enix" and year_of_release IS NOT NULL
GROUP BY year_of_release
ORDER BY year_of_release
```

| Year | North_America_Sales | Europe_Sales | Japan_Sales | Other_Sales | Global_Sales |
|---|---|---|---|---|---|
| 2003 | 168000 | 86000 | 150000 | 13000 | 416000 |

| Year | North_America_Sales | Europe_Sales | Japan_Sales | Other_Sales | Global_Sales |
|---|---|---|---|---|---|
| 2004 | 12000 | 9000 | 0 | 3000 | 25000 |
| 2005 | 233000 | 68000 | 185000 | 20000 | 506000 |
| 2006 | 363000 | 52000 | 373000 | 200000 | 988000 |
| 2007 | 273000 | 93000 | 168000 | 62000 | 597000 |
| 2008 | 76000 | 70000 | 109000 | 37000 | 291000 |
| 2009 | 245000 | 141000 | 227000 | 60000 | 672000 |
| 2010 | 294000 | 159000 | 166000 | 64000 | 683000 |
| 2011 | 135000 | 114000 | 218000 | 37000 | 505000 |
| 2012 | 88000 | 26000 | 130000 | 10000 | 254000 |
| 2013 | 46000 | 28000 | 87000 | 13000 | 175000 |
| 2015 | 2000 | 7000 | 7000 | 1000 | 17000 |
| 2016 | 12000 | 20000 | 16000 | 7000 | 54000 |




## COMPLETE CODE BELOW

```
# %%
import numpy as np

import altair as alt

import datadotworld as dw

import pandas as pd

import opendatasets as od

# %%
# dw.query('https://data.world/mhoangvslev/steam-games-dataset',
# '''
# SELECT
# FROMobject has no attribute 'properties'
# WHERE
# ORDER BY
# LIMIT
# '''
# )


# %% [markdown]
# ## Find the top 10 Cities with the most amount of game developers

# %%
q1 =dw.query('https://data.world/mhoangvslev/steam-games-dataset',
'''
SELECT DISTINCT city AS City, country AS Country, COUNT(developer) as Developer_Count
FROM developper
GROUP BY city
ORDER BY COUNT(developer) DESC
LIMIT 10
'''
)

# %%
(print(q1.dataframe
.to_markdown(index=None)))

# %%
q1_chart = (alt.Chart(q1.dataframe, title = "Top 10 Cities With The Most Games Developers")
.mark_bar()
.encode(
    x = "Developer_Count",
    y = "City")
.properties(
    width = 800,
    height = 400
))

q1_chart
```

```
# %%
q1_chart.save("q1_chart.png")

# %% [markdown]
# ## Find the Top 5 devleopers with the highest amount of sales (global Sales)

# %%
q2 = dw.query('https://data.world/mhoangvslev/steam-games-dataset',
'''
SELECT d.developer AS Developer, sum(v.global_sales)*100000 AS Global_Sales
FROM vgsales v
JOIN developper d
ON  v.developer = d.developer
GROUP BY d.developer
ORDER BY Global_Sales DESC
LIMIT 10
'''
)

# %%
(print(q2.dataframe
.to_markdown(index=None)))

# %%
q2_chart = (alt.Chart(q2.dataframe, title = "Top 5 Developers by Global Sales")
.mark_bar()
.encode(
    x="Developer",
    y="Global_Sales")
.properties(
    width = 800,
    height = 400
))

q2_chart

# %%
q2_chart.save("q2_chart.png")

# %% [markdown]
# ## Breakdown the amount of games Nintendo has developed by Genre

# %%
q3 = dw.query('https://data.world/mhoangvslev/steam-games-dataset',
'''
SELECT genre AS Genre, COUNT(GENRE) AS Number_of_Games_Made
FROM vgsales
WHERE developer = "Nintendo"
GROUP BY genre
ORDER BY COUNT(GENRE) ASC
'''
)
```

```
# %%
(print(q3.dataframe
.to_markdown(index=None)))

# %%
q3_chart = (alt.Chart(q3.dataframe, title = "Games Made by Nintendo Broken Down By Genre")
.mark_bar()
.encode(
    x="Number_of_Games_Made",
    y="Genre")
.properties(
    width = 800,
    height = 400
))

q3_chart

# %%
q3_chart.save("q3_chart.png")

# %% [markdown]
# ## Create a tabel that shows the sales of Square Enix (developer) by year, include all of the sale

# %%
q4 = dw.query('https://data.world/mhoangvslev/steam-games-dataset',
'''
SELECT year_of_release as Year,
    SUM(na_sales) AS North_America_Sales,
    SUM(eu_sales) AS Europe_Sales ,
    SUM(jp_sales) AS Japan_Sales,
    SUM(other_sales) AS Other_Sales,
    SUM(global_sales) AS Global_Sales
FROM vgsales
WHERE developer = "Square Enix" and year_of_release IS NOT NULL
GROUP BY year_of_release
ORDER BY year_of_release
'''
)

# %%
(print(q4.dataframe
.to_markdown(index=None)))

# %%
q4_chart_na = (alt.Chart(q4.dataframe, title = "Sales of Sqaure Enix Games by Area")
.mark_line(line = True)
.encode(
    x=alt.X('Year:O'),
    y=alt.Y('North_America_Sales', title=""))
.properties(
    width = 800,
    height = 400
```

```
))

q4_chart_eu = (alt.Chart(q4.dataframe, title = "Sales of Sqaure Enix Games by Area")
.mark_line(line = True, color='blue')
.encode(
    x=alt.X('Year:O'),
    y=alt.Y('Europe_Sales', title=""))
)

q4_chart_jp = (alt.Chart(q4.dataframe, title = "Sales of Sqaure Enix Games by Area")
.mark_line(line = True, color='geen')
.encode(
    x=alt.X('Year:O'),
    y=alt.Y('Japan_Sales', title=""))
)

q4_chart_other = (alt.Chart(q4.dataframe, title = "Sales of Sqaure Enix Games by Area")
.mark_line(line = True, color='purple')
.encode(
    x=alt.X('Year:O'),
    y=alt.Y('Other_Sales', title=""))
)

q4_chart_global = (alt.Chart(q4.dataframe, title = "Sales of Sqaure Enix Games by Area")
.mark_line(line = True, color='orange')
.encode(
    x=alt.X('Year:O'),
    y=alt.Y('Global_Sales', title=""))
)

# %%
q4_chart = q4_chart_na+q4_chart_eu+q4_chart_jp+q4_chart_other+q4_chart_global

q4_chart

# %%
q4_chart.save("q4_chart.png")
```