




Game Project

(Section 1)

임규현
AI 부트캠프 10기
2021.12.13



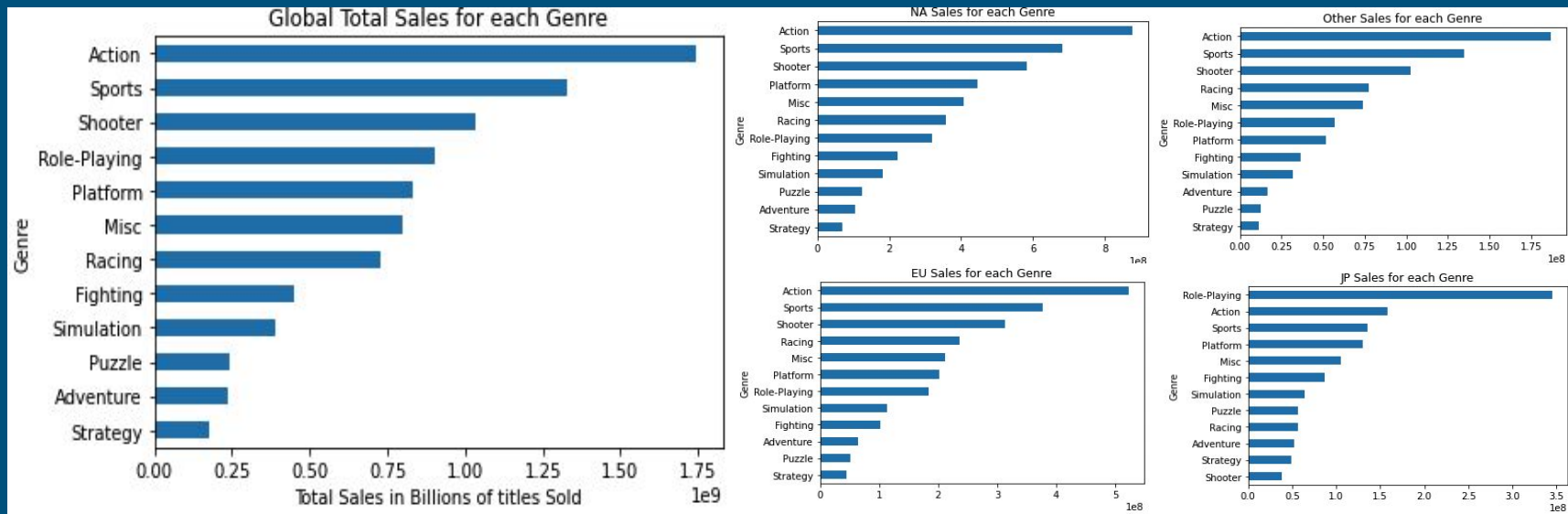
문제 제시 (Problem Statement)

- Q. 다음 분기 어떤 게임을 설계해야 할까?
 - 지역에 따라서 선호하는 게임 장르가 다른가?
 - 연도별 게임 트렌드?
 - 출고량이 높은 게임에 대한 분석

1. 지역에 따라서 선호하는 게임 장르가 다른가?

- 1

- 각 지역별 장르 출고량 데이터 정리
 - 일본이 조금 다름 다름 (RPG, Action, Sports, Platform, Misc)



1. 지역에 따라서 선호하는 게임 장르가 다른가?

- 2

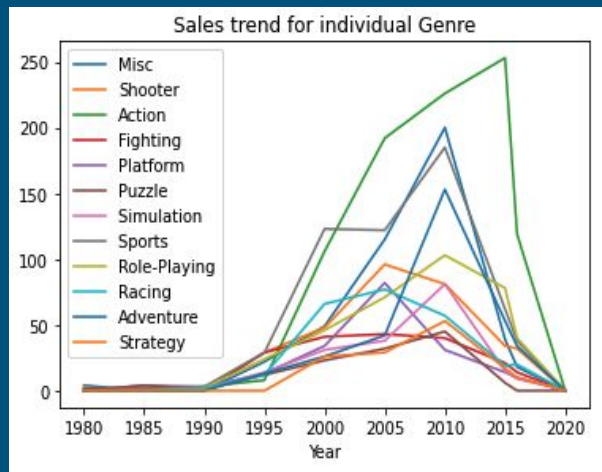
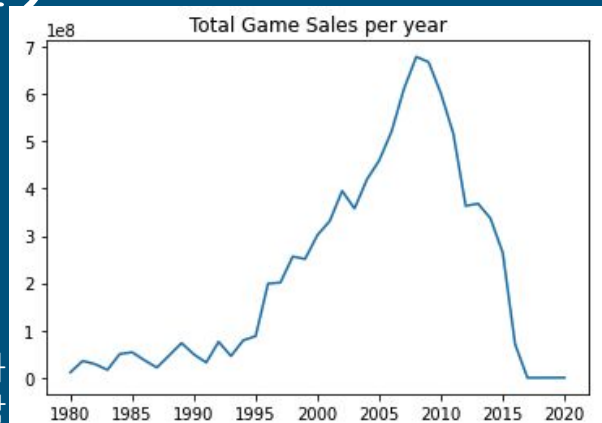
- 지역별 출고량의 분포?
- F-stat 가설검정 - 여러 그룹들의 분포
 - # H0: 지역별 각 장르 마다 출고량의 분포에 차이가 없다
 - # Ha: 지역별 각 장르 마다 출고량의 분포에 차이가 있다
 - #P-value 가 유의 수준 0.05 p-value 보다 작기 때문에 귀무가설을 버리고 대립가설을 채택
 - 결론: 지역별 출고량의 분포에 차이가 있다
 - 지역에 따라 선호하는 게임 장르가 다르다 / 게임 선호에 대한 분포가 다르다 라는 결론

Genre	Action	Adventure	Fighting	Misc	Platform	Puzzle	Racing	Role-Playing	Shooter	Simulation	Sports	Strategy
NA_Sales	8.748100e+08	105660000.0	223200000.0	408050000.0	446440000.0	123350000.0	357190000.0	319540000.0	5.819200e+08	182960000.0	6.805200e+08	68700000.0
EU_Sales	5.221600e+08	64050000.0	101220000.0	212360000.0	201610000.0	50730000.0	237100000.0	183580000.0	3.130500e+08	113350000.0	3.764800e+08	45320000.0
JP_Sales	1.579000e+08	51950000.0	87340000.0	105270000.0	130650000.0	56870000.0	56400000.0	344950000.0	3.828000e+07	63700000.0	1.353700e+08	49460000.0
Other_Sales	1.864500e+08	16810000.0	36640000.0	74040000.0	51530000.0	12520000.0	77100000.0	56960000.0	1.026500e+08	31480000.0	1.346700e+08	11360000.0
Total	1.741320e+09	238470000.0	448400000.0	799720000.0	830230000.0	243470000.0	727790000.0	905030000.0	1.035900e+09	391490000.0	1.327040e+09	174840000.0

F_onewayResult(statistic=9.055483346936597, pvalue=8.756119956631937e-05)

2. 연도별 게임의 트렌드가 있을까?

- 연도별 게임의 정리
- 전반적인 트렌드 - 연도별 총 판매량
 - 총판매량 피크 후 급속한 하락 (2010년)
 - 90~2000년대 초반까지는 스포츠 게임이 시장의 강자
 - 액션 게임의 경우 거의 2000년대 들어서야 시장의 강자로 자리잡음
 - 2003년 이후로는 액션 게임이 시장을 지배하기 시작함
- 최근 몇 년 동안은 데이터가 별로 없음?
 - 출고량 급감의 외부적인 요인?
 - 데이터셋 자체의 문제인지?
 - 최근 몇 년 간의 트렌드를 파악하는데 어려움이 있음



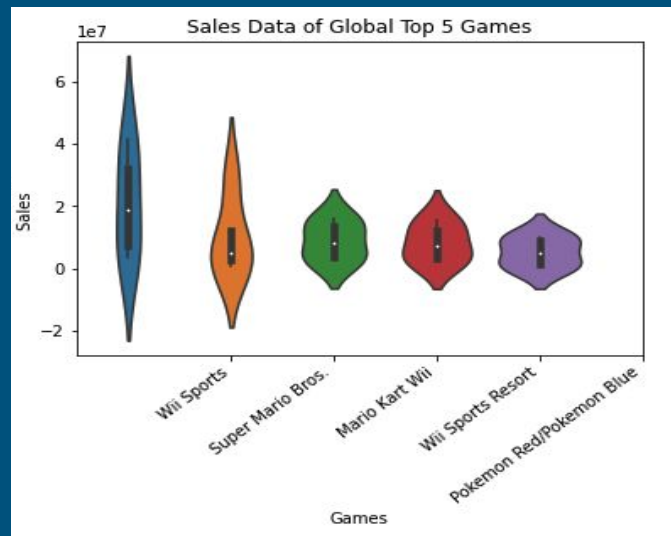
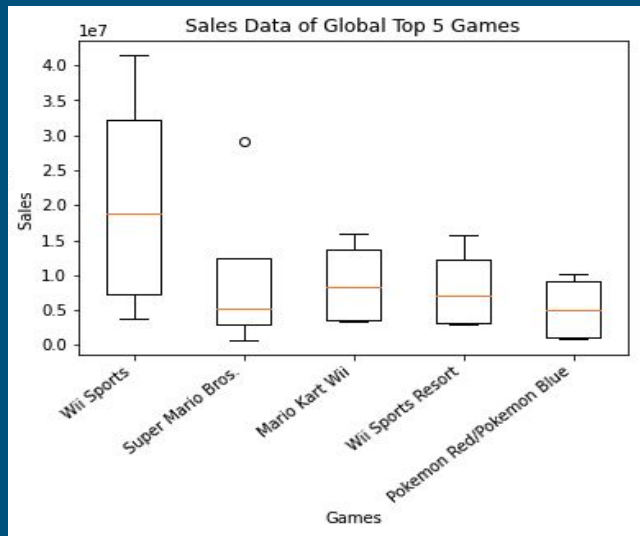
3. 출고량이 높은 게임에 대한 분석 - 1

- 어떤 게임들?
- 글로벌 총 판매량 순 Top 10 -> Nintendo
 - 장르- 플랫폼, 스포츠, 레이싱, RPG 등

	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Total	Average
Name										
Wii Sports	Wii	2006.0	Sports	Nintendo	41490000.0	29020000.0	3770000.0	8460000.0	82740000.0	20685000.0
Super Mario Bros.	NES	1985.0	Platform	Nintendo	29080000.0	3580000.0	6810000.0	770000.0	40240000.0	10060000.0
Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15850000.0	12880000.0	3790000.0	3310000.0	35830000.0	8957500.0
Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15750000.0	11010000.0	3280000.0	2960000.0	33000000.0	8250000.0
Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11270000.0	8890000.0	10220000.0	1000000.0	31380000.0	7845000.0
Tetris	GB	1989.0	Puzzle	Nintendo	23200000.0	2260000.0	4220000.0	580000.0	30260000.0	7565000.0
New Super Mario Bros.	DS	2006.0	Platform	Nintendo	11380000.0	9230000.0	6500000.0	2900000.0	30010000.0	7502500.0
Wii Play	Wii	2006.0	Misc	Nintendo	14030000.0	9200000.0	2930000.0	2850000.0	29010000.0	7252500.0
New Super Mario Bros. Wii	Wii	2009.0	Platform	Nintendo	14590000.0	7060000.0	4700000.0	2260000.0	28610000.0	7152500.0
Duck Hunt	NES	1984.0	Shooter	Nintendo	26930000.0	630000.0	280000.0	470000.0	28310000.0	7077500.0

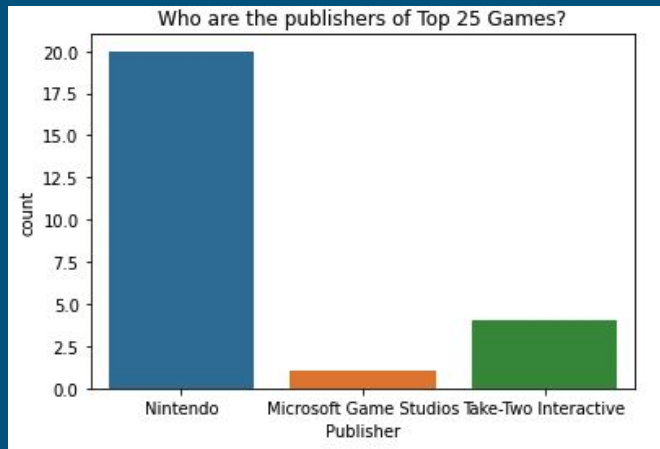
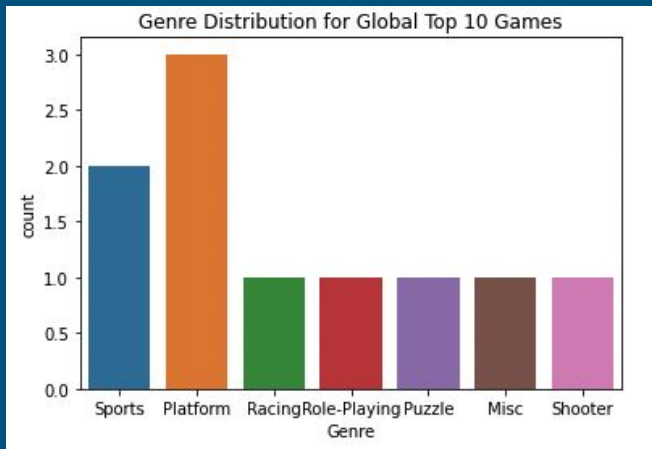
3. 출고량이 높은 게임에 대한 분석 - 2

- Global Top 5 Games 에 대한 시각화
 - Box , Violin



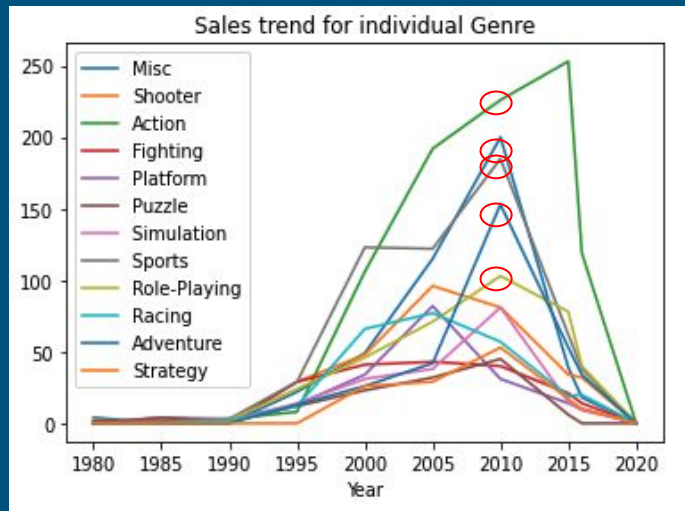
3. 출고량이 높은 게임에 대한 분석 - 3

- 출고량 Top 10 게임들의 장르 분포도
 - 플랫폼, 스포츠 순
 - 액션? (10~25위권)
- 출고량 Top 25 게임들의 Publisher
 - Nintendo가 압도적

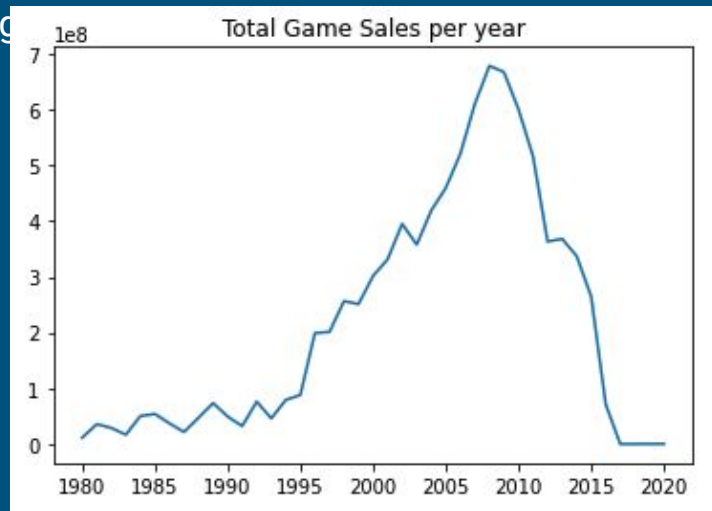


4. 그렇다면 어떤 게임을 설계? - 1

- 여러가지 방법이 있겠지만 시장의 평균적인 퍼포먼스를 따라가는 방법을 선택
 - 회사에 대한 몇 가지 가정
 - 우리의 클라이언트는 시장의 키플레이어 X (아직 게임 출시 X)
 - 일단 시장에 진출하는 것이 목표 (안정적?)



e-Playing



4. 그렇다면 어떤 게임을 설계? - 2

- 설계하고자 하는 개별 장르와 **Total Sales**의 상관 관계에 대해서 생각
 - Misc, Sports, Action, Adventure, Role-Playing
- 상관계수 - 어떤 변수가 **Total Sales**와 연관성이 큰가?
 - Misc가 가장 높다
- 'Misc' 카테고리에 해당하는 게임을 설계하는 걸 추천

```
# MISC 와 Total Sales의 상관관계
cov2 = np.cov(x,y)[0,1].      #공분산
cor2 = np.corrcoef(x,y)[0,1]  #상관관계
cor2
```

0.9995501000485616

```
x = [408050000.0, 212360000.0, 105270000.0, 74040000.0, 799720000.0] #MISC
x1 = [680520000.0, 376480000.0, 135370000.0, 134670000.0, 13270400000.0] #Sports
x2 = [874810000.0, 522160000.0, 157900000.0, 186450000.0, 1741320000.0] #Action
x3 = [105660000.0, 64050000.0, 51950000.0, 16810000.0, 238470000.0] #Adventure
y = [4392950000, 2434130000, 1291020000, 797750000, 8915850000] #Total Sales
```

4. 그렇다면 어떤 게임을 설계? - 3

```
# Sports 와 Total Sales의 상관관계
cov2 = np.cov(x1,y)[0,1]
cor2 = np.corrcoef(x1,y)[0,1]
cor2
```

0.9230521640531281

```
# Adventure 와 Total Sales의 상관관계
cov2 = np.cov(x3,y)[0,1]
cor2 = np.corrcoef(x3,y)[0,1]
cor2
```

0.992616099676239

```
# Action 과 Total Sales의 상관관계
cov2 = np.cov(x2,y)[0,1]
cor2 = np.corrcoef(x2,y)[0,1]
cor2
```

0.9964427305562199

```
# Role-Playing 과 Total Sales의 상관관계
cov2 = np.cov(x4,y)[0,1]      #공분산
cor2 = np.corrcoef(x4,y)[0,1] #상관관계
cor2
```

0.9240916641508188

The End

이상으로 프로젝트 1 발표를 마치겠습니다. 감사합니다 :)

추후 제안 사항 (Next Steps)

- 데이터셋 정비/점검
 - 최근 몇 년간에 대한 데이터를 구할 수 있는지?
-