

Section1 : Project

다음 분기 게임 설계를 위한 데이터 분석



CONTENTS



#1

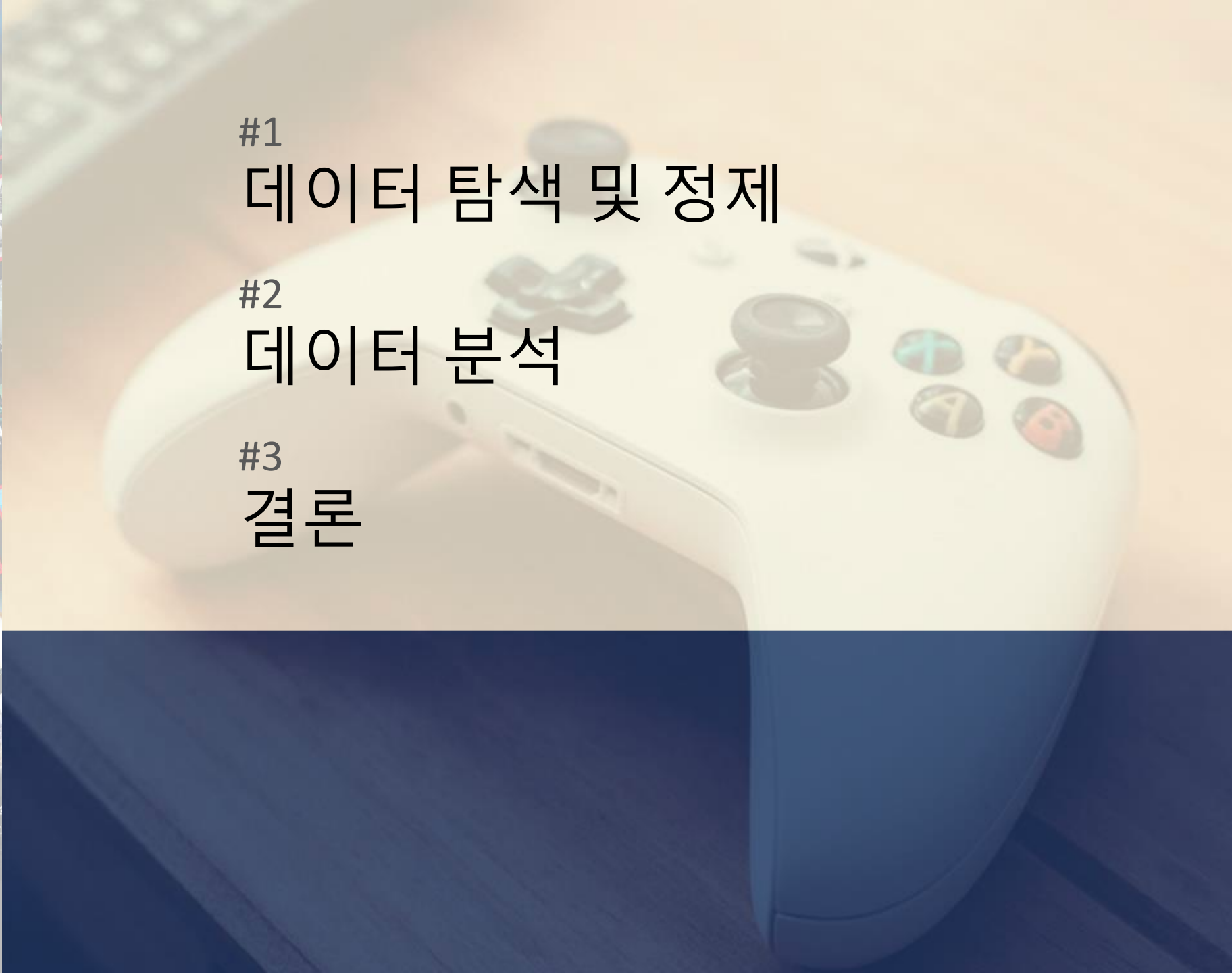
데이터 탐색 및 정제

#2

데이터 분석

#3

결론



1. 데이터 탐색 및 정제

a. 품질의 문제

	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales
8973	Hayate no Gotoku! Ojousama Produce Daisakusen ...	DS	2008.0	Simulation	Konami Digital Entertainment	0	0	0.04	0
13635	Vanguard Bandits	PS	1998.0	Strategy	Human Entertainment	0.04	0.03	0	0
5654	Jillian Michaels Fitness Ultimatum 2010	DS	2009.0	Sports	Majesco Entertainment	0.07	0	0	0.01
14743	LEGO The Lord of the Rings	DS	2012.0	Action	Warner Bros. Interactive Entertainment	0	0.14	0	0.02
12209	Family Fest Presents Circus Games	Wii	2008.0	Misc	Ubisoft	0.1	0.29	0	0.05
13218	Time Commando	PS	1996.0	Action	Adeline Software	0.1	0.07	0	0.01
13085	CandyLand / Chutes & Ladders / Memory	GBA	2005.0	Misc	Zoo Digital Publishing	0.15	0.05	0	0
14292	Tom Clancy's Rainbow Six: Critical Hour	XB	NaN	Shooter	Unknown	0.04	0.01	0	0
1376	Doraemon 2: Nobita no Toizurando Daibouken	SNES	1993.0	Role-Playing	Epoch	0	0	0.05	0
11334	Super Fire ProWrestling X	SNES	1995	Fighting	Human Entertainment	0	0	0.28	0

- Unnamed: 0은 필요 없는 column
- 지역별 판매량의 기본 단위는 million / M과 K는 불필요
- 출시연도 column의 이상데이터 처리
- 결측값 처리

1. 데이터 탐색 및 정제

a. 품질의 문제

Unnamed: 0		Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales
31	32	Rock Band	X360	NaN	Misc	Electronic Arts	1.93	0.34	0	0.21
50	51	Ape Escape: On the Loose	PSP	2005.0	NaN	Sony Computer Entertainment	0.52	0.01	0.12	0.05
109	110	Yu-Gi-Oh! 5D's Wheelie Breakers (JP sales)	Wii	NaN	Racing	Unknown	0	0	0.02	0
272	273	UK Truck Simulator	PC	2010.0	Simulation	NaN	0	0.03	0	0
273	274	Jewel Link Chronicles: Mountains of Madness	DS	NaN	Puzzle	Avanquest	0	0.06	0	0.01
...
16425	16426	Yoostar on MTV	X360	NaN	Misc	Unknown	0.06	0	0	0
16431	16432	Cars: Mater-National Championship	X360	2007.0	NaN	THQ	0.13	0	0	0.01
16432	16433	Bejeweled 3	X360	NaN	Puzzle	Unknown	0.13	0	0	0.01
16555	16556	Apache: Air Assault	PC	2010.0	NaN	Just Flight	0	0.03	0	0.01
16587	16588	Need for Speed Underground 2	XB	2004.0	NaN	Electronic Arts	1.38	0.8	0	0.1

357 rows × 10 columns

- Unnamed: 0은 필요 없는 column
- 지역별 판매량의 기본 단위는 million / M과 K는 불필요
- 출시연도 column의 이상데이터 처리
- 결측값 처리

1. 데이터 탐색 및 정제

b. 구조적 문제

Unnamed: 0		Name	Platform	Year	Genre	Publisher
0	1	Candace Kane's Candy Factory	DS	2008.0	Action	Destineer
1	2	The Munchables	Wii	2009.0	Action	Namco Bandai Games
2	3	Otome wa Oanesama Boku ni Koi Shiteru Portable	PSP	2010.0	Adventure	Alchemist
3	4	Deal or No Deal: Special Edition	DS	2010.0	Misc	Zoo Games
4	5	Ben 10 Ultimate Alien: Cosmic Destruction	PS3	2010.0	Platform	D3Publisher

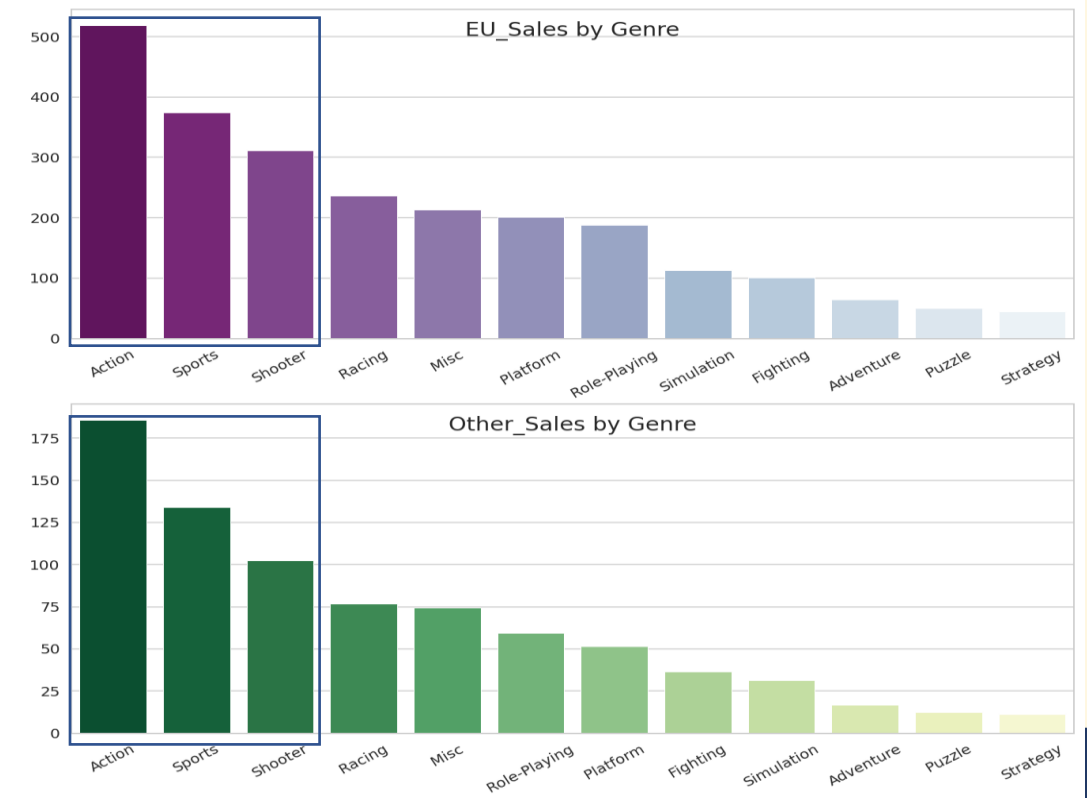
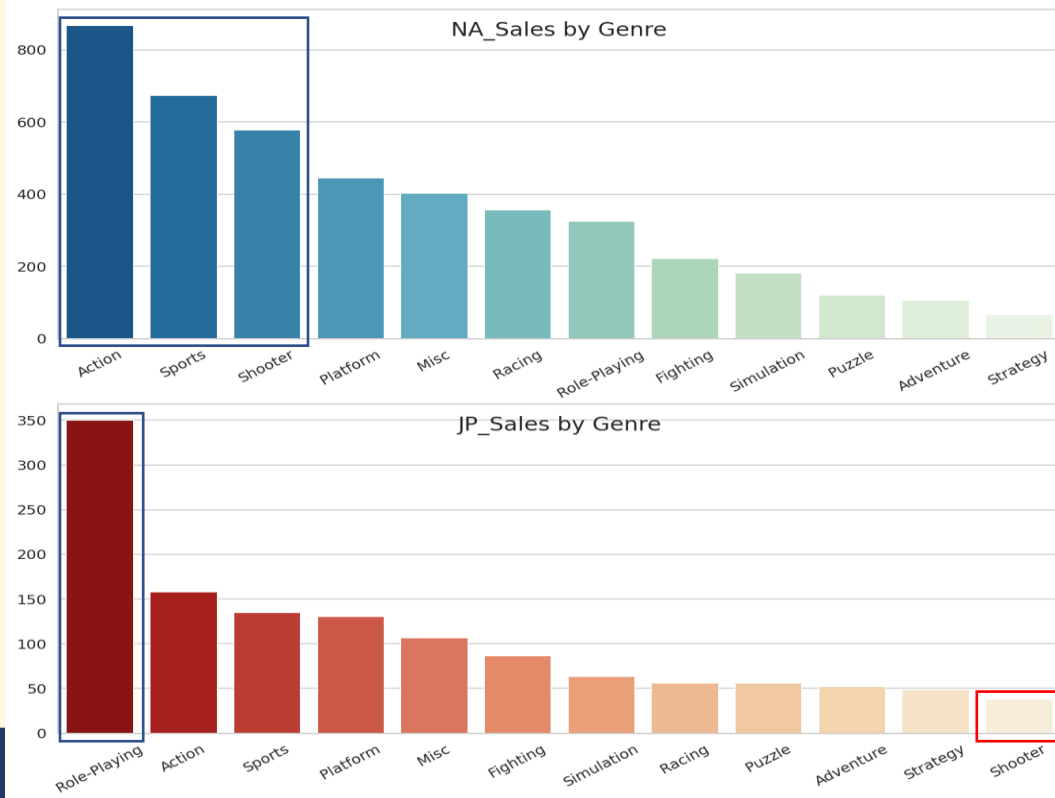
Country	Sales
NA	0.04
EU	0.00
JP	0.00
Other	0.00
NA	0.17
EU	0.00
JP	0.00
Other	0.01
NA	0.00
EU	0.00

- Tidy Data가 되기 위해서 지역과 출고량이 각각의 column이 되도록 구성

2. 데이터 분석

a. 지역별 게임 장르 선호도

출고량 합계 기준

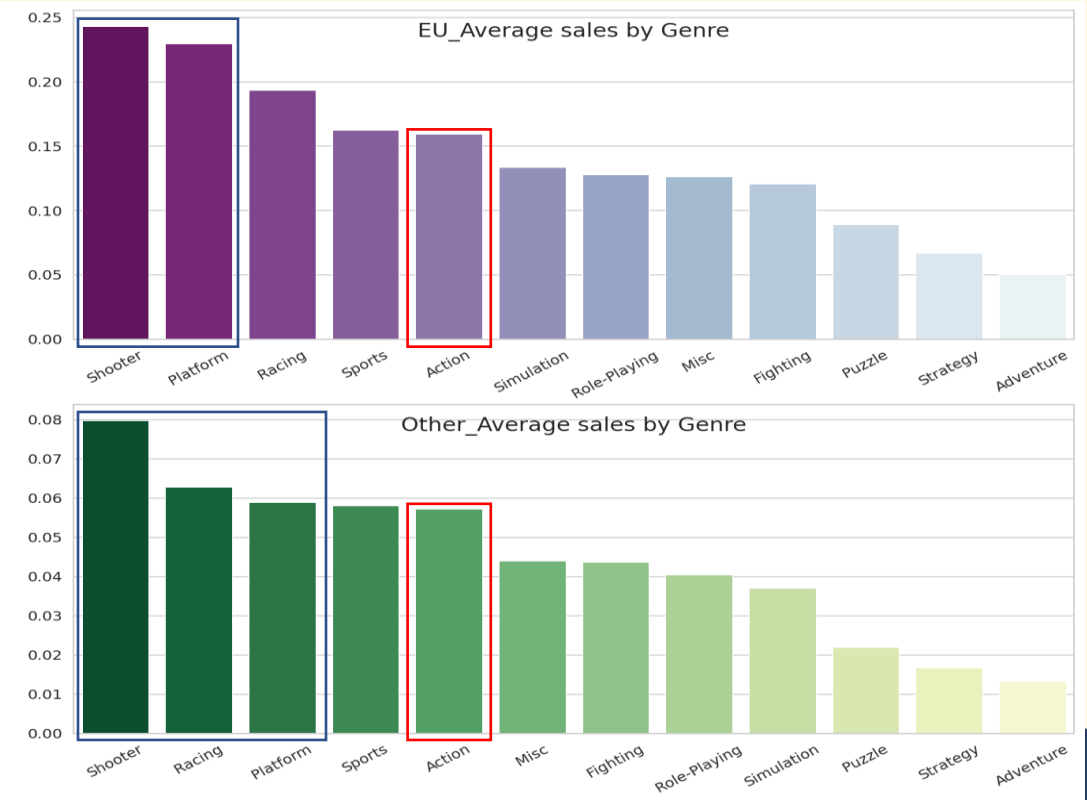
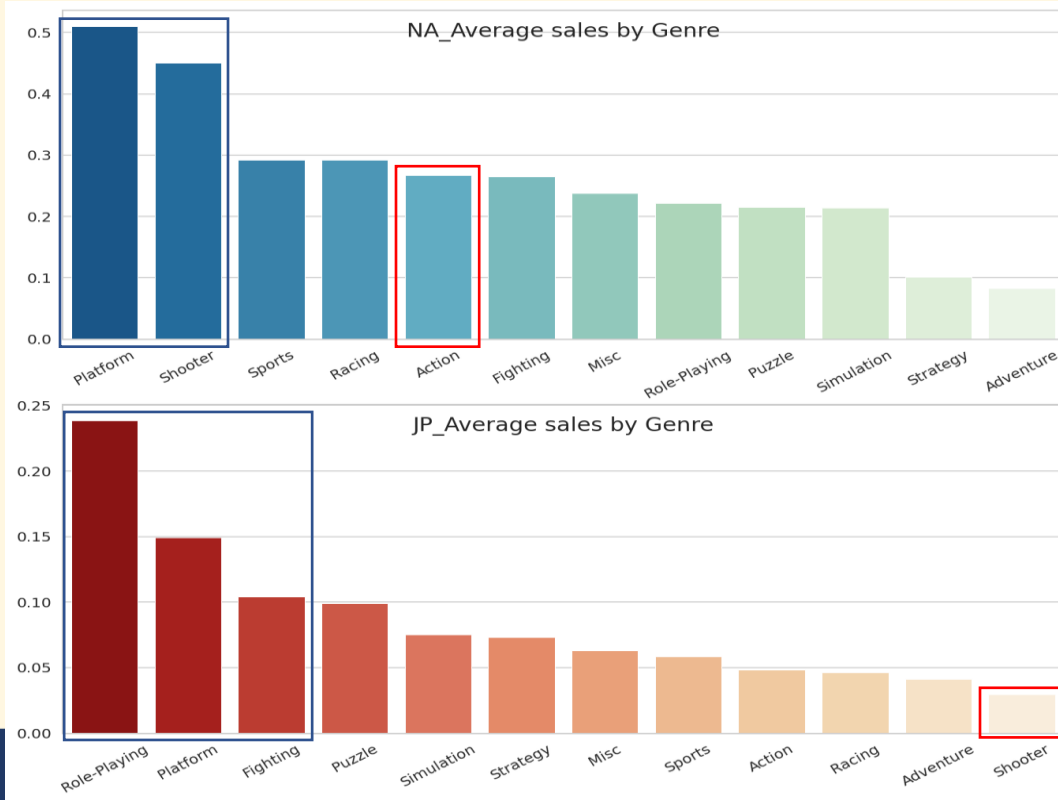


- Action과 Sports 장르의 높은 출고량
- 일본은 다른 지역과 다른 양상
- 평균 출고량은 Platform과 Shooter 장르가 강세
- 일본은 합계와 평균 모두 Role-Playing이 압도적

2. 데이터 분석

a. 지역별 게임 장르 선호도

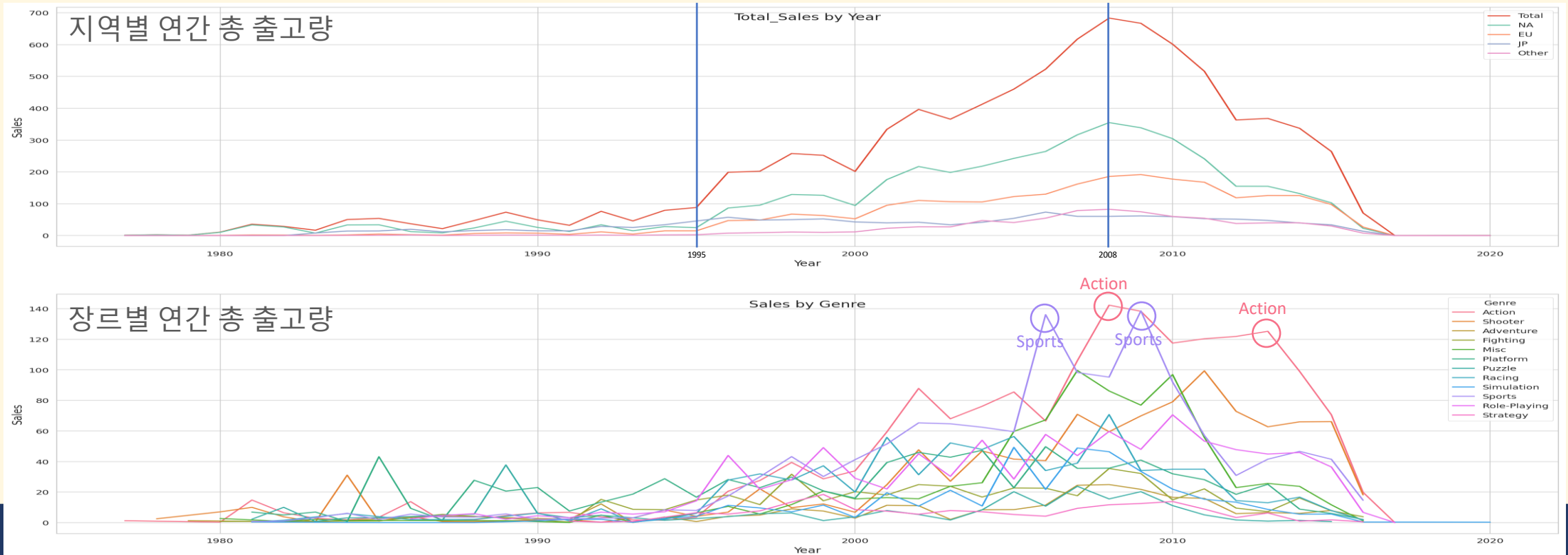
출고량 평균 기준



- Action과 Sports 장르의 높은 출고량
- 일본은 다른 지역과 다른 양상
- 평균 출고량은 Platform과 Shooter 장르가 강세
- 일본은 합계와 평균 모두 Role-Playing이 압도적

2. 데이터 분석

b. 연도별 게임 장르 선호도

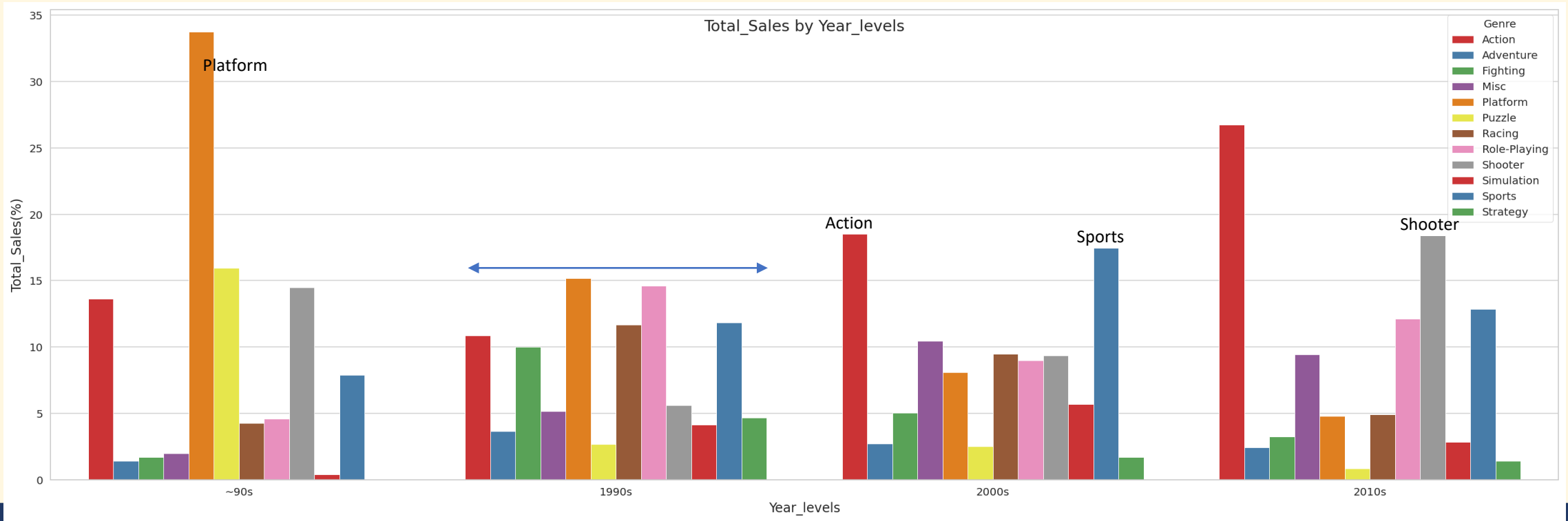


- 전세계 게임 시장은 1995년까지 꾸준한 상승세와 1996~2008년 까지 가파른 상승세
- 2008년 기점으로 게임시장은 지속적인 하락세
- 장르별 판매량은 Action과 Sports의 강세

2. 데이터 분석

b. 연도별 게임 장르 선호도

시대별 장르 비율



90년대 이전 : Platform 장르가 압도적

1990년대 : 다양한 장르가 경쟁

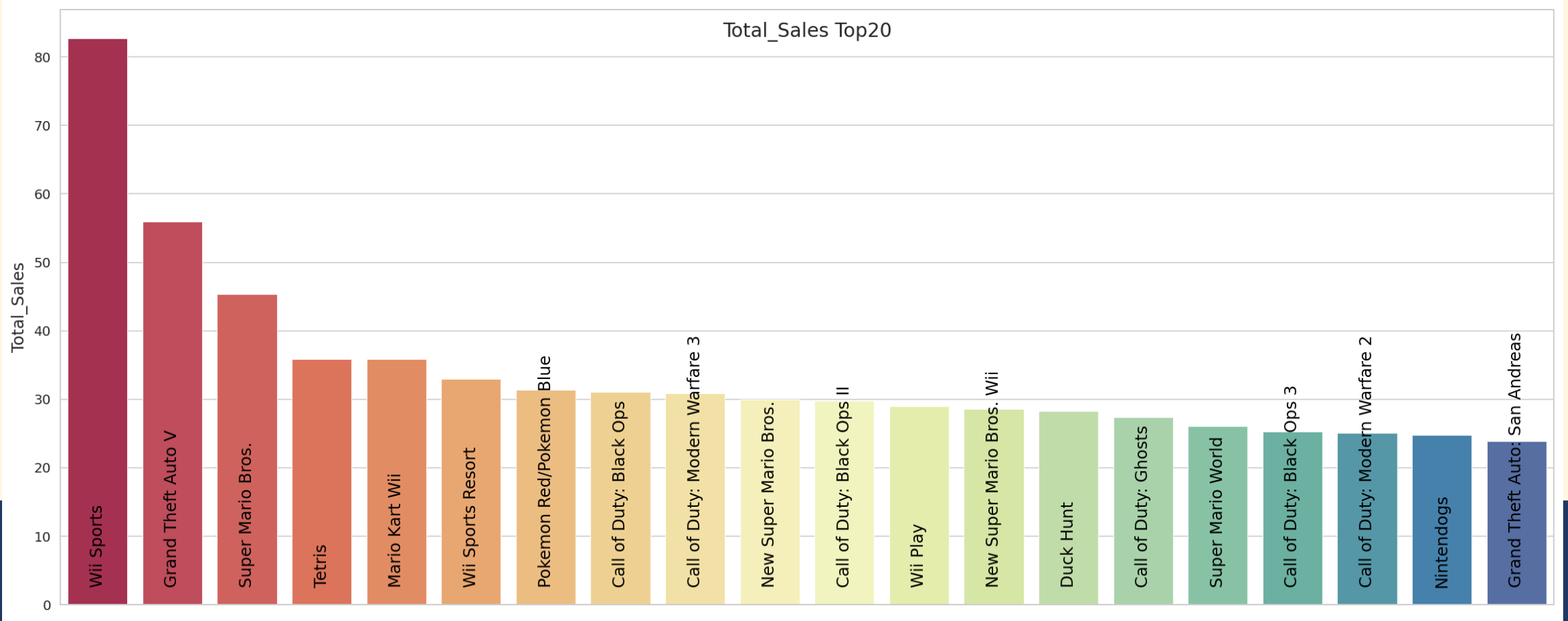
2000년대 : Action과 Sports의 성장

2010년대 : Shooter가 새롭게 떠오름

2. 데이터 분석

c. 높은 출고량의 게임

게임 출고량 Top20

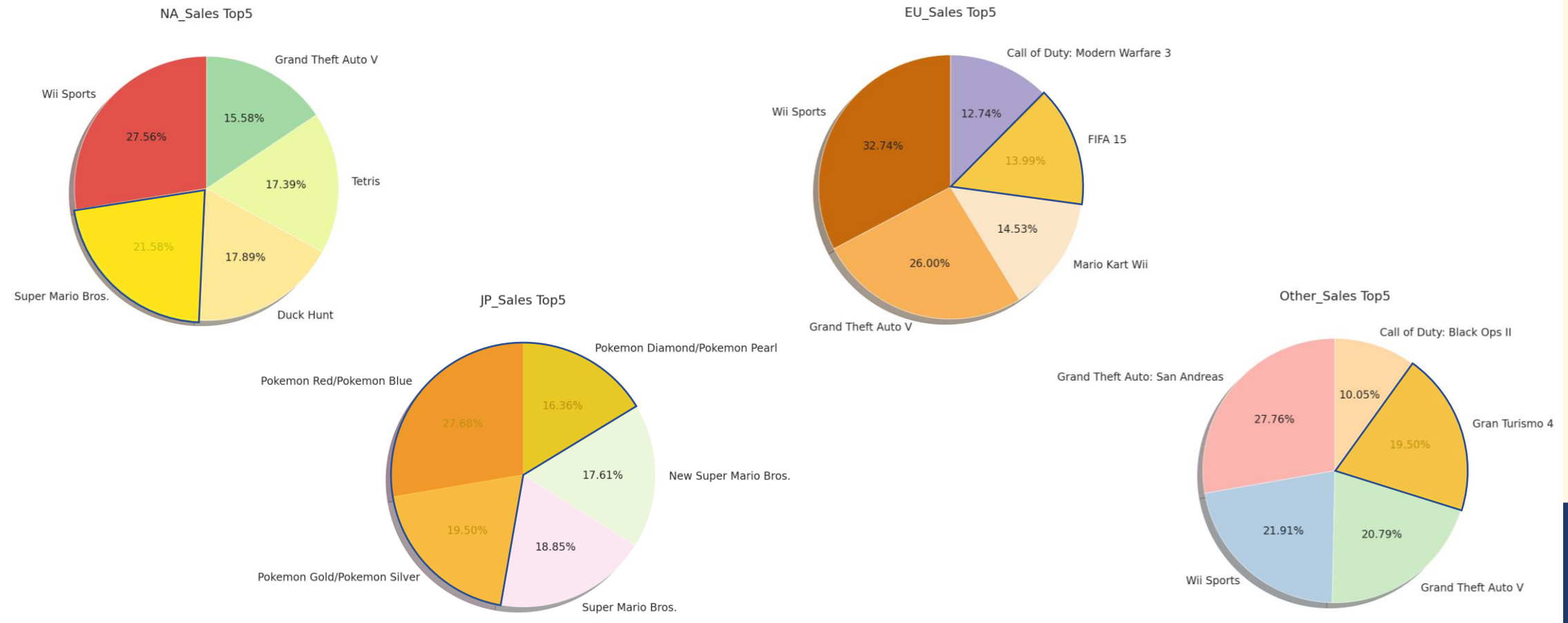


- Wii Sports의 높은 출고량
- Action 장르는 GTA 시리즈 뿐

2. 데이터 분석

c. 높은 출고량의 게임

지역별 Top5 게임

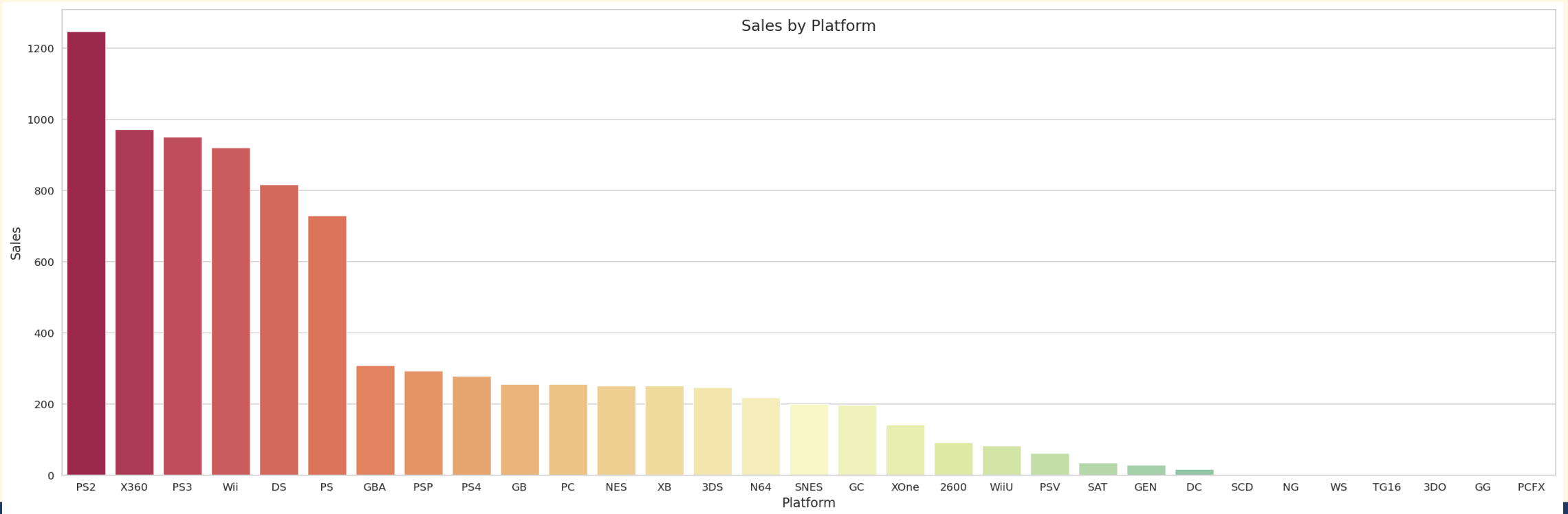


- 북미의 Super Mario Bros.(Platform), 유럽의 FIFA15(Sports), 기타 지역의 Gran Turismo 4(Racing)
- 일본은 Pokemon Series가 강세

2. 데이터 분석

d. 높은 출고량의 플랫폼

플랫폼 별 출고량

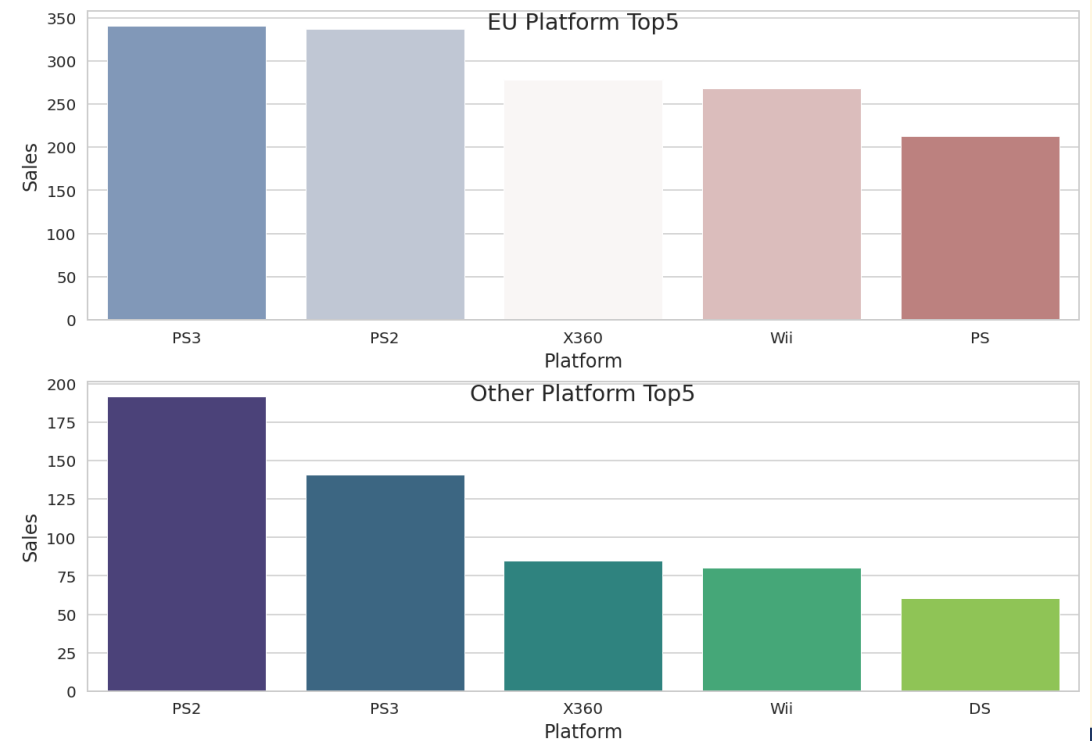
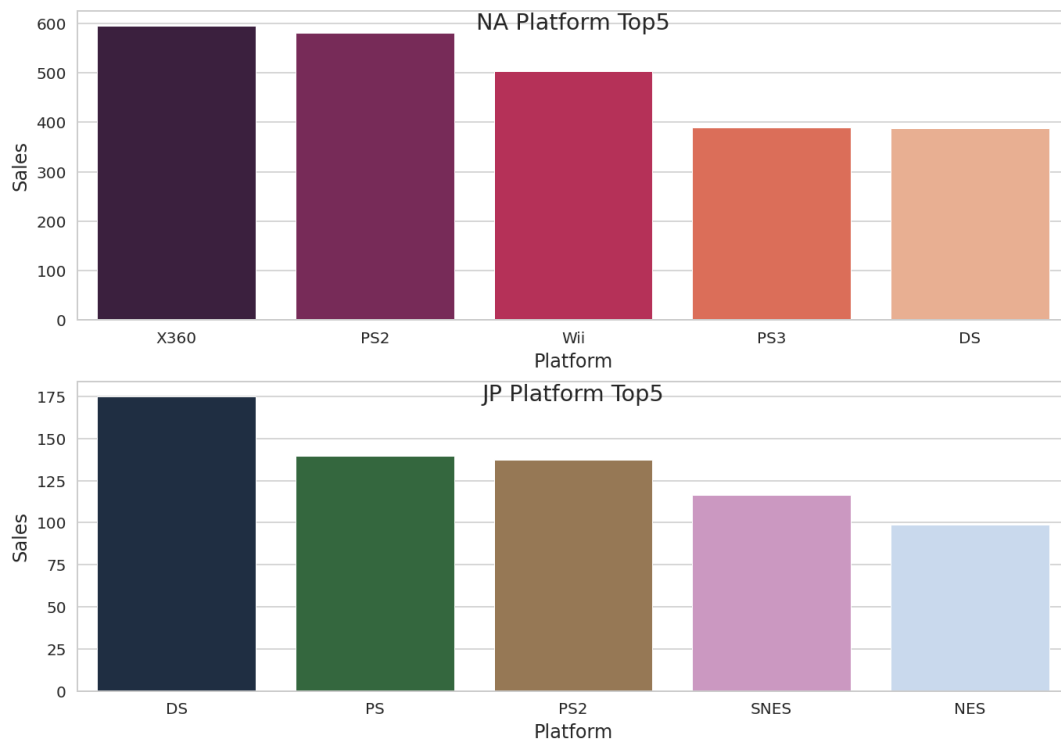


- 1위는 PS2 / X360과 PS3, Wii가 각축
- 북미는 X360이 우위
- 유럽은 트렌디한 모습을 보임
- 일본은 DS가 압도적 / X360은 5위권 밖

2. 데이터 분석

d. 높은 출고량의 플랫폼

지역별 플랫폼 Top5



- 1위는 PS2 / X360과 PS3, Wii가 각축
- 북미는 X360이 우위
- 유럽은 트렌디한 모습을 보임
- 일본은 DS가 압도적 / X360은 5위권 밖

2. 데이터 분석

e. 이원분산분석(two-way Anova)

df_anova			
	Sales	Year_levels	Genre
0	1.25	~90s	Action
1	2.53	~90s	Shooter
2	1.30	~90s	Adventure
3	0.34	~90s	Action
4	0.77	~90s	Fighting
...
386	14.63	2010s	Sports
387	0.52	2010s	Strategy
388	0.01	2010s	Action
389	0.04	2010s	Role-Playing
390	0.29	2010s	Simulation

391 rows x 3 columns

```
[93] # statsmodels를 활용한 이원분산분석
from statsmodels.formula.api import ols
from statsmodels.stats.anova import anova_lm

model = ols('Sales ~ C(Year_levels) * C(Genre)', df_anova).fit()
anova_lm(model)
```

	df	sum_sq	mean_sq	F	PR(>F)
C(Year_levels)	3.0	71302.939069	23767.646356	86.527045	9.797292e-42
C(Genre)	11.0	68376.232088	6216.021099	22.629668	1.055535e-34
C(Year_levels):C(Genre)	33.0	61508.192768	1863.884629	6.785545	3.960076e-22
Residual	344.0	94491.500860	274.684596	NaN	NaN

- 귀무가설 : 시대와 장르에 따른 출고량의 평균은 같다.
- 대립가설 : 시대와 장르에 따른 출고량의 평균은 다르다.
- 종속변수 : Sales
- 독립변수 : Year_levels, Genre
- 독립변수가 두개 이상이므로 이원분산분석(two-way Anova) 사용

2. 데이터 분석

e. 이원분산분석(two-way Anova)

df_anova

	Sales	Year_levels	Genre
0	1.25	~90s	Action
1	2.53	~90s	Shooter
2	1.30	~90s	Adventure
3	0.34	~90s	Action
4	0.77	~90s	Fighting
...
386	14.63	2010s	Sports
387	0.52	2010s	Strategy
388	0.01	2010s	Action
389	0.04	2010s	Role-Playing
390	0.29	2010s	Simulation

391 rows x 3 columns

```
[93] # statsmodels를 활용한 이원분산분석
from statsmodels.formula.api import ols
from statsmodels.stats.anova import anova_lm

model = ols('Sales ~ C(Year_levels) + C(Genre)', df_anova).fit()
anova_lm(model)
```

	df	sum_sq	mean_sq	F	PR(>F)
C(Year_levels)	3.0	71302.939069	23767.646356	86.527045	9.797292e-42
C(Genre)	11.0	68376.232088	6216.021099	22.629668	1.055535e-34
C(Year_levels):C(Genre)	33.0	61508.192768	1863.884629	6.785545	3.960076e-22
Residual	344.0	94491.500860	274.684596	NaN	NaN

P-value

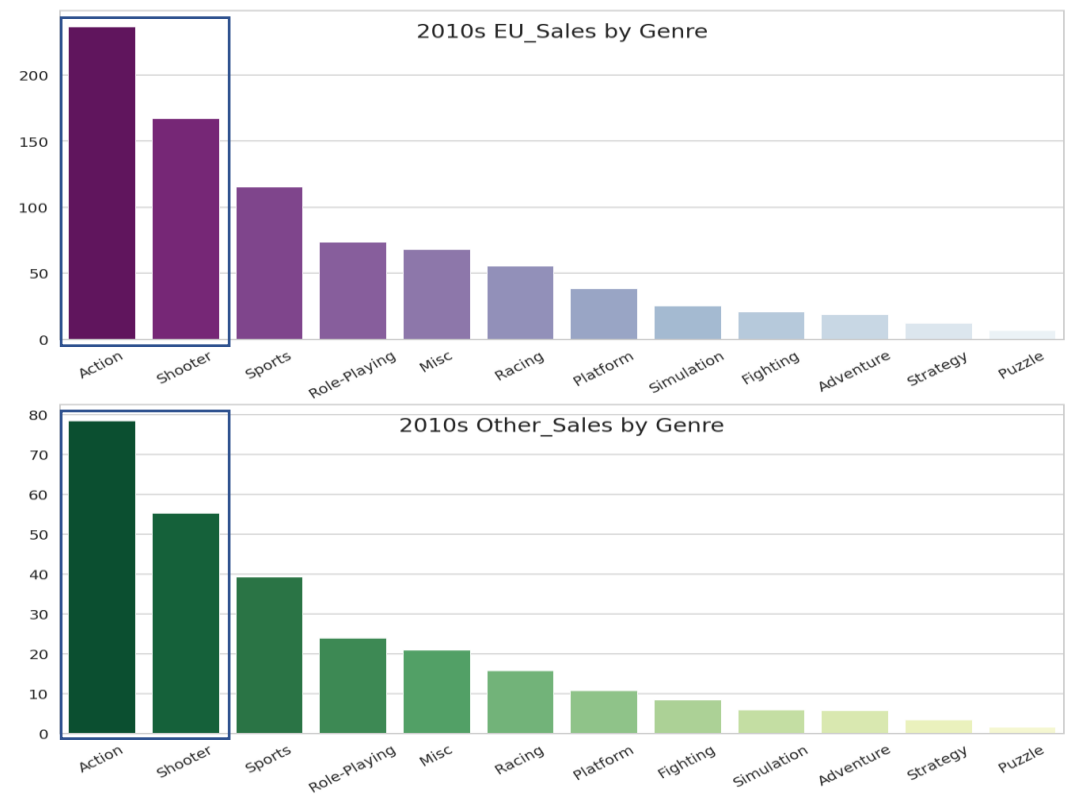
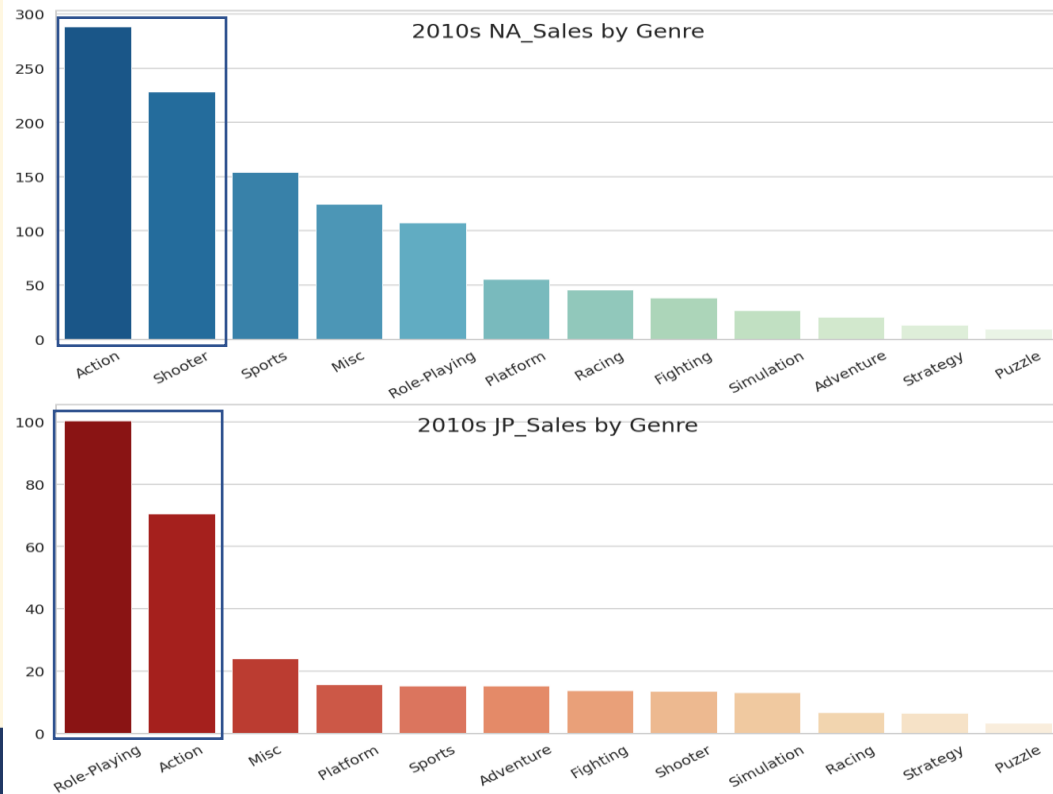
관찰된 데이터가 귀무가설과 양립하는 정도를 0에서 1 사이의 수치로 표현한 것
p-value가 작을수록 그 정도가 약함

- Year_levels : $p < 0.05$ 로 유의미. Year_levels에 따라 평균에 차이가 발생
- Genre : $p < 0.05$ 로 유의미. Genre에 따라 평균에 차이 발생
- Year_levels:Genre : $p < 0.05$ 로 유의미. 상호작용 효과 있음

2. 데이터 분석

f. 2010년 이후 게임 데이터 분석

지역별 장르 선호도(출고량 합계 기준)

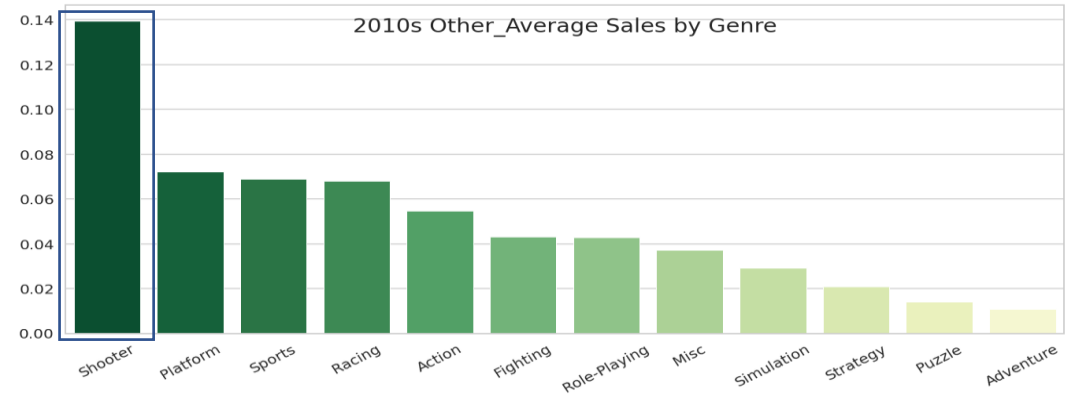
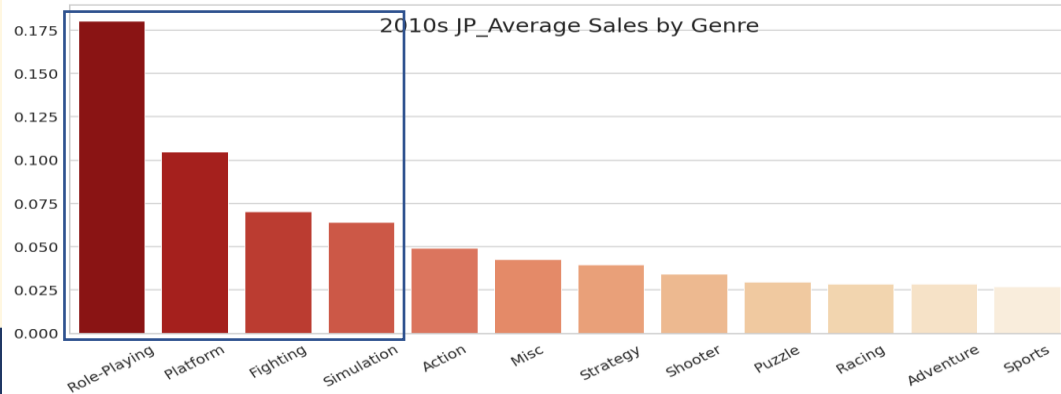
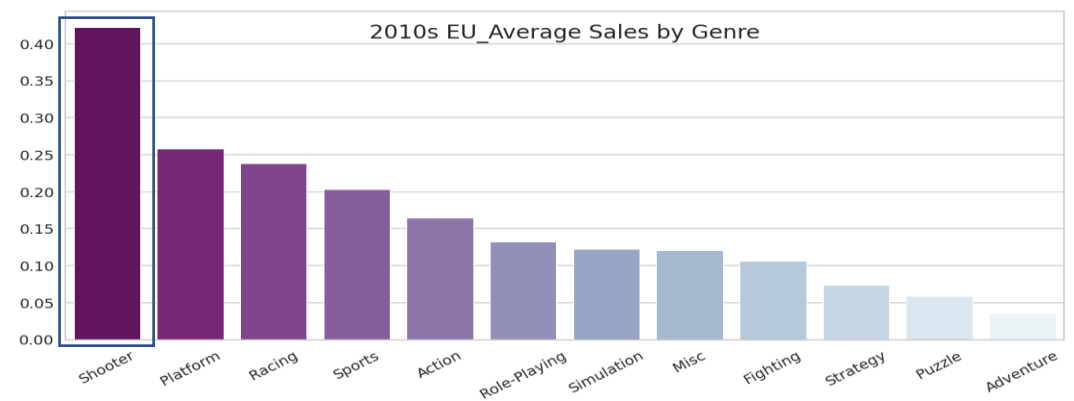
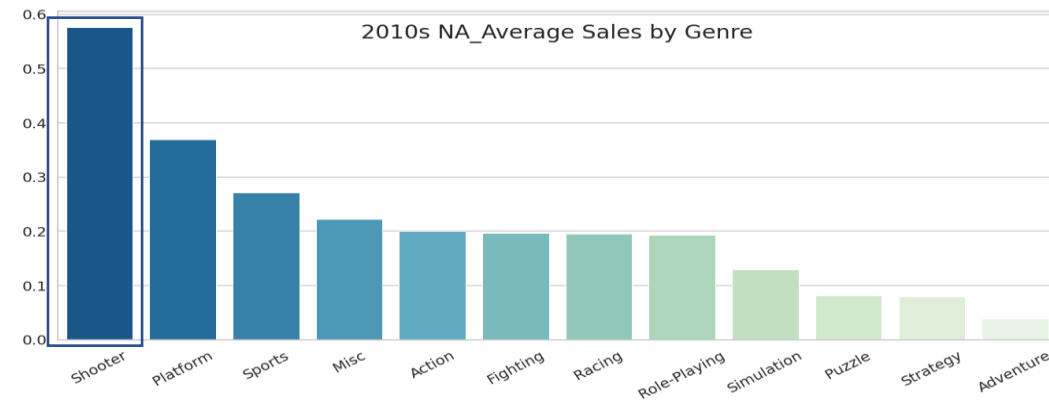


- 출고량 합계에서 Action 장르의 강세는 유효
- 일본은 여전히 Role-Playing이 가장 높고 Action 장르의 출고량이 유의미
- 출고량 평균에선 Shooter가 압도적
- 일본은 평균 출고량도 Role-Playing이 가장 높고 Fighting, Simulation이 선전중

2. 데이터 분석

f. 2010년 이후 게임 데이터 분석

지역별 장르 선호도(출고량 평균 기준)

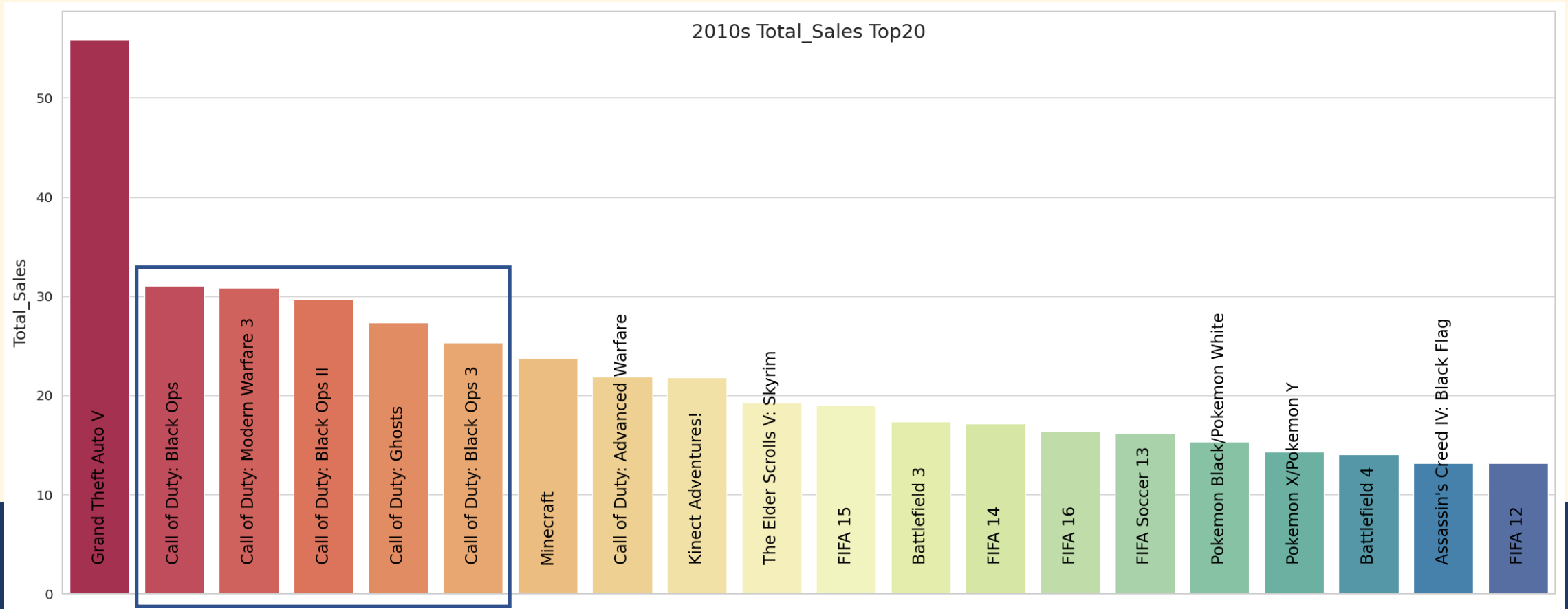


- 출고량 합계에서 Action 장르의 강세는 유효
- 일본은 여전히 Role-Playing이 가장 높고 Action 장르의 출고량이 유의미
- 출고량 평균에선 Shooter가 압도적
- 일본은 평균 출고량도 Role-Playing이 가장 높고 Fighting, Simulation이 선전중

2. 데이터 분석

f. 2010년 이후 게임 데이터 분석

2010년 이후 게임 Top20

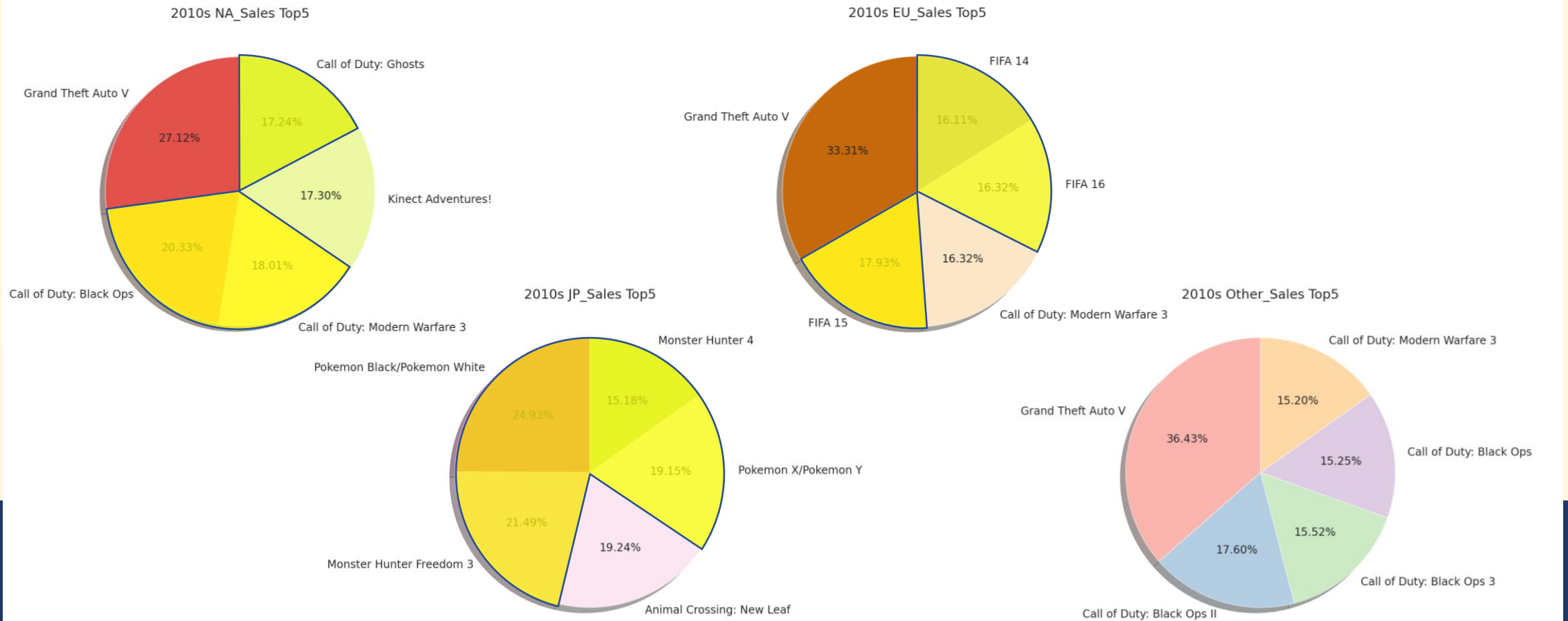


- 출고량 1위 GTA5입니다.
- Call of Duty, FIFA, Battlefield, Pokemon과 같이 Series로 제작된 게임의 강세

2. 데이터 분석

f. 2010년 이후 게임 데이터 분석

2010년 이후 지역별 Top5 게임

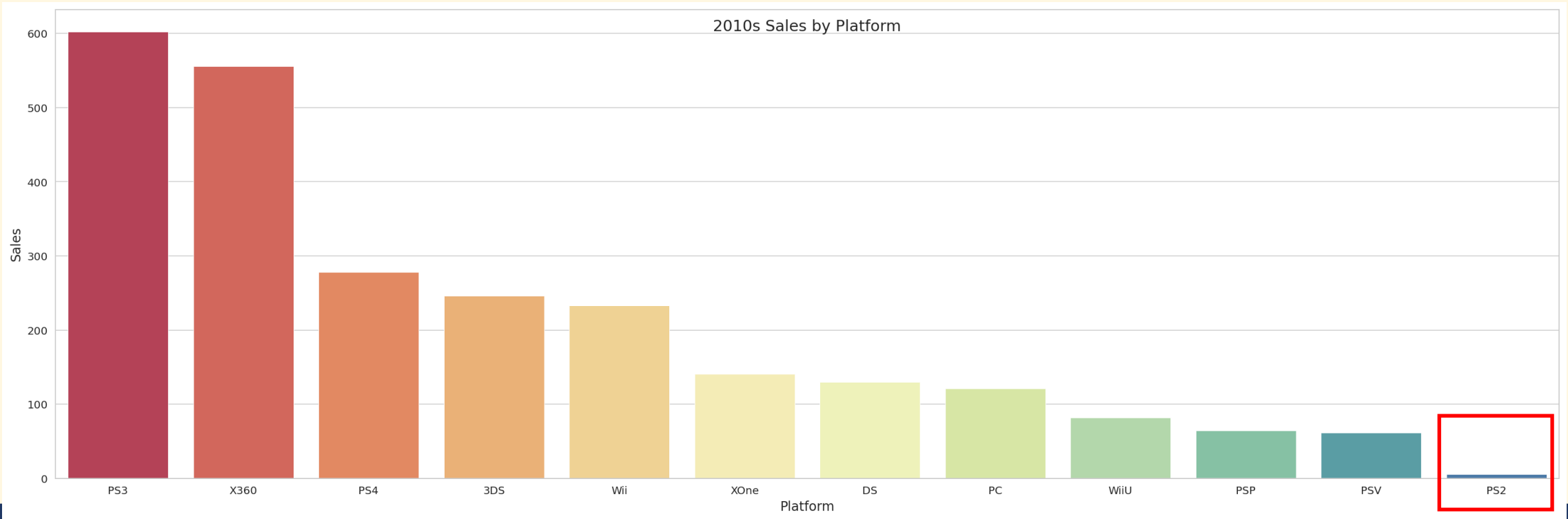


- 지역별로 선호 게임 차이가 큼
-북미의 Call of Duty, 유럽의 FIFA, 일본의 Pokemon과 Monster Hunter-
- 일본 순위 안에 GTA5가 없음

2. 데이터 분석

f. 2010년 이후 게임 데이터 분석

2010년 이후 플랫폼 별 출고량

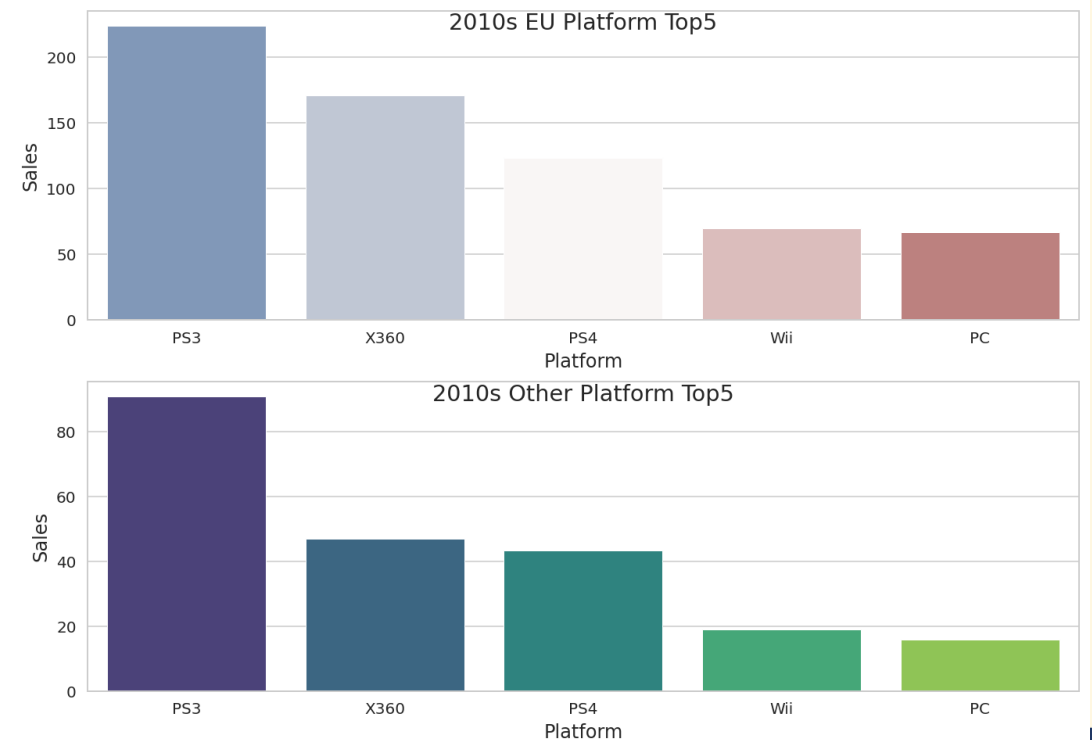
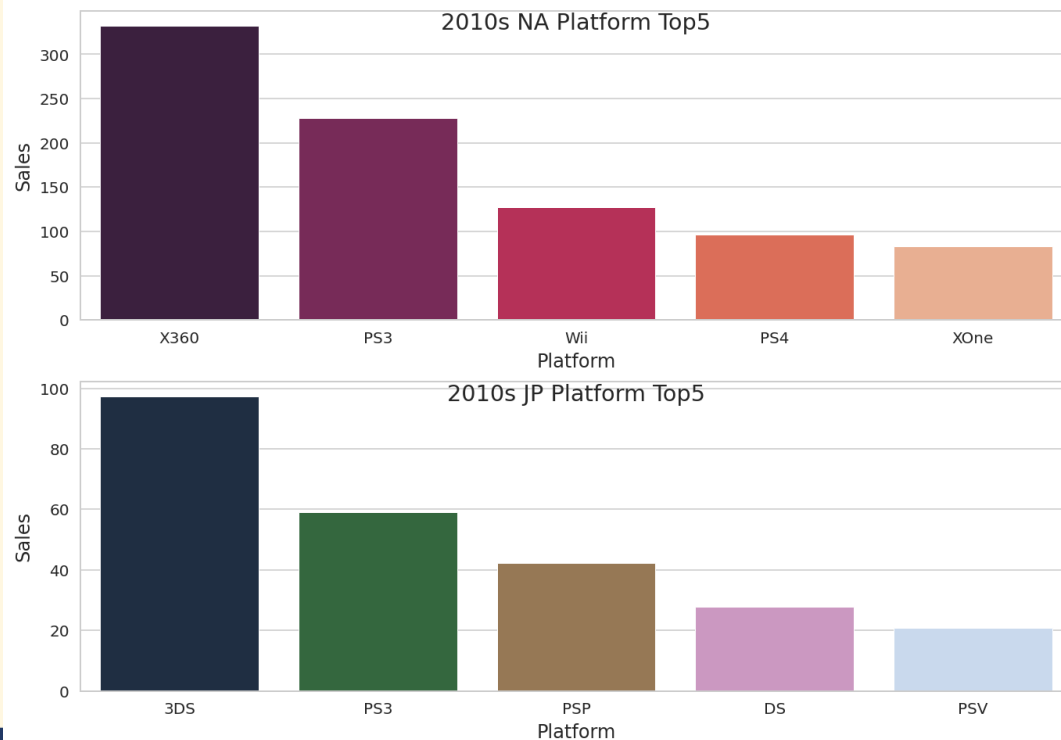


- 최하위의 PS2와 PS4의 상승세
- PS3와 X360의 치열한 1위 싸움
- 여전한 북미의 X360 사랑
- 유럽은 여전히 트렌디한 모습
- 일본은 DS의 자리를 3DS가 대체

2. 데이터 분석

f. 2010년 이후 게임 데이터 분석

2010년 이후 지역별 플랫폼 Top5

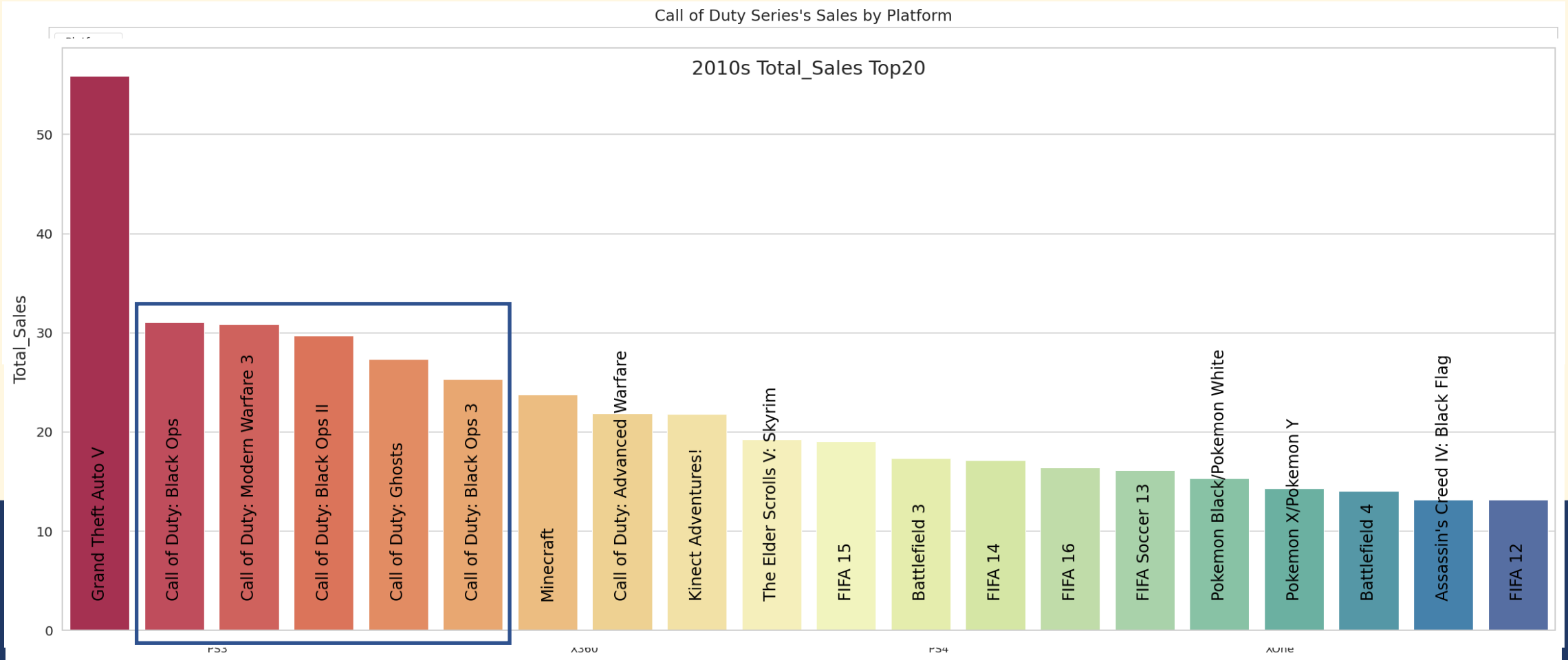


- 최하위의 PS2와 PS4의 상승세
- PS3와 X360의 치열한 1위 싸움
- 여전한 북미의 X360 사랑
- 유럽은 여전히 트렌디한 모습
- 일본은 DS의 자리를 3DS가 대체

2. 데이터 분석

f. 2010년 이후 게임 데이터 분석

Call of Duty Series

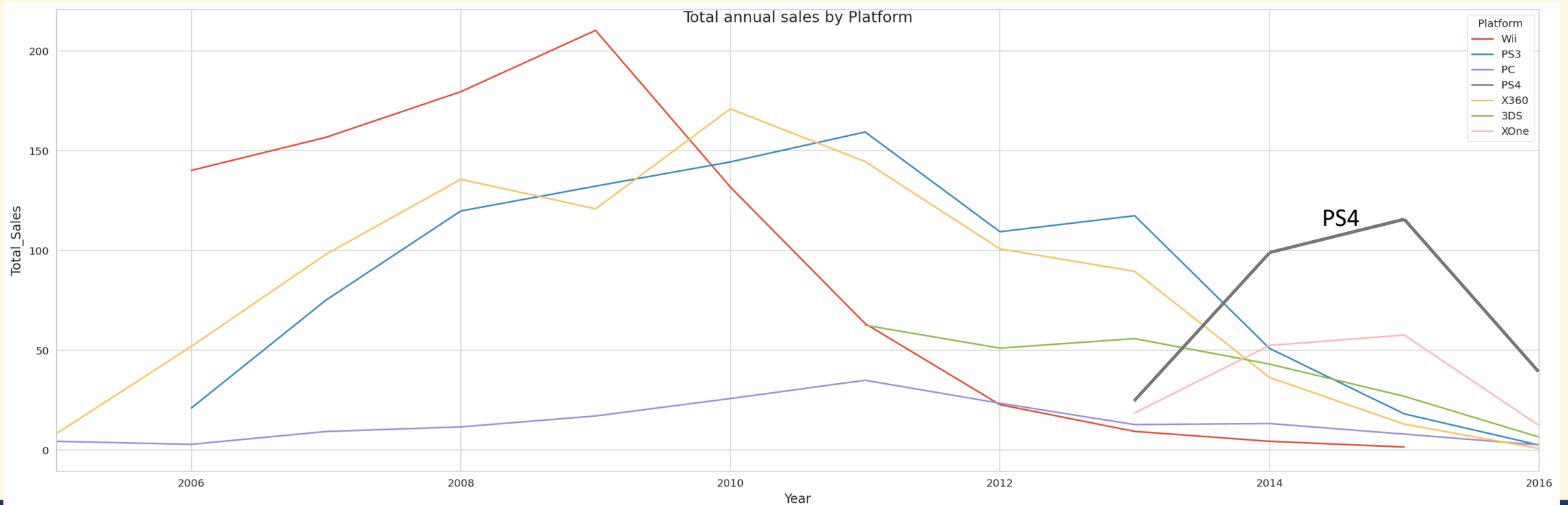


- 시대가 변함에 따라 PS3와 X360에서 PS4와 Xone으로 넘어가는 흐름
- PS3와 X360는 균형을 이루었으나 PS4는 Xone과 격차를 벌임

2. 데이터 분석

f. 2010년 이후 게임 데이터 분석

2010년 이후 플랫폼별 연간 총 출고량



- 게임 시장의 규모가 감소 추세
- PS4의 약진
- PC게임 시장의 안정적인 흐름

3. 결론

- Shooter 장르를 PS4와 X One 멀티플랫폼으로 설계
- 일본 시장까지 고려한다면 Platform 장르와 PS4 플랫폼을 대상으로 삼는 것도 좋은 선택
- 장르가 시장에서 강력한 경쟁을 하기 위해 세계 시장 및 매력적인 IP를 참조



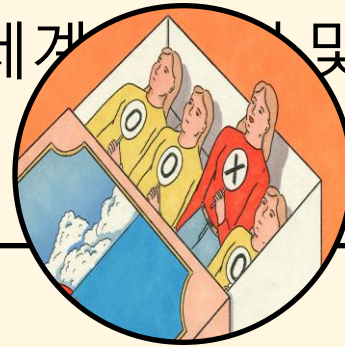
Action의 함정, Shooter에 주목

- Action장르의 강세는 출시한 게임수의 영향
- Shooter 장르의 선호도가 가장 높음
- 일본까지 고려하면 Platform 장르도 고려 대상



시대에 맞는 플랫폼

- 최다출고량의 PS2는 2010년 이후 최하위
- PS3와 X360의 강세
- PS4와 XOne의 상승세
- 안정적인 PC게임 시장 흐름



결이 다른 일본

- GTA5의 저조한 성적
- 압도적인 Role-Playing
- Shooter는 하위권
- 3DS가 DS의 자리를 대체
- 낮은 점유율의 Xbox Series



게임 IP의 강력함

- 출고량 순위 상위권을 독식한 Series 게임
- 북미의 Call of Duty
- 유럽의 FIFA
- 일본의 Pokemon과 Monster Hunter

감사합니다