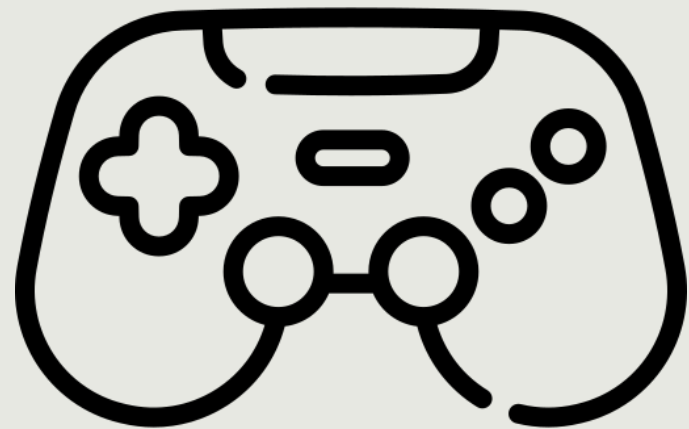




AI 부트캠프 SECTION 01 PROJECT

AI 부트캠프 10기 박민경



목차

00 데이터 전처리

01 지역별 게임 장르 선호도에 대한 분석

02 연도별 트렌드 분석

03 출고량이 높은 게임에 대한 분석

04 다음 분기에 설계해야 할 게임

✓ 결측치: drop하여 처리

-> 년도에 해당하는 컬럼에 있는 결측치는 평균을 내거나 특정한 년도 데이터로 대체할 수 없다고 판단해 drop으로 처리

✓ 지역별 출고량 데이터 단위 통일

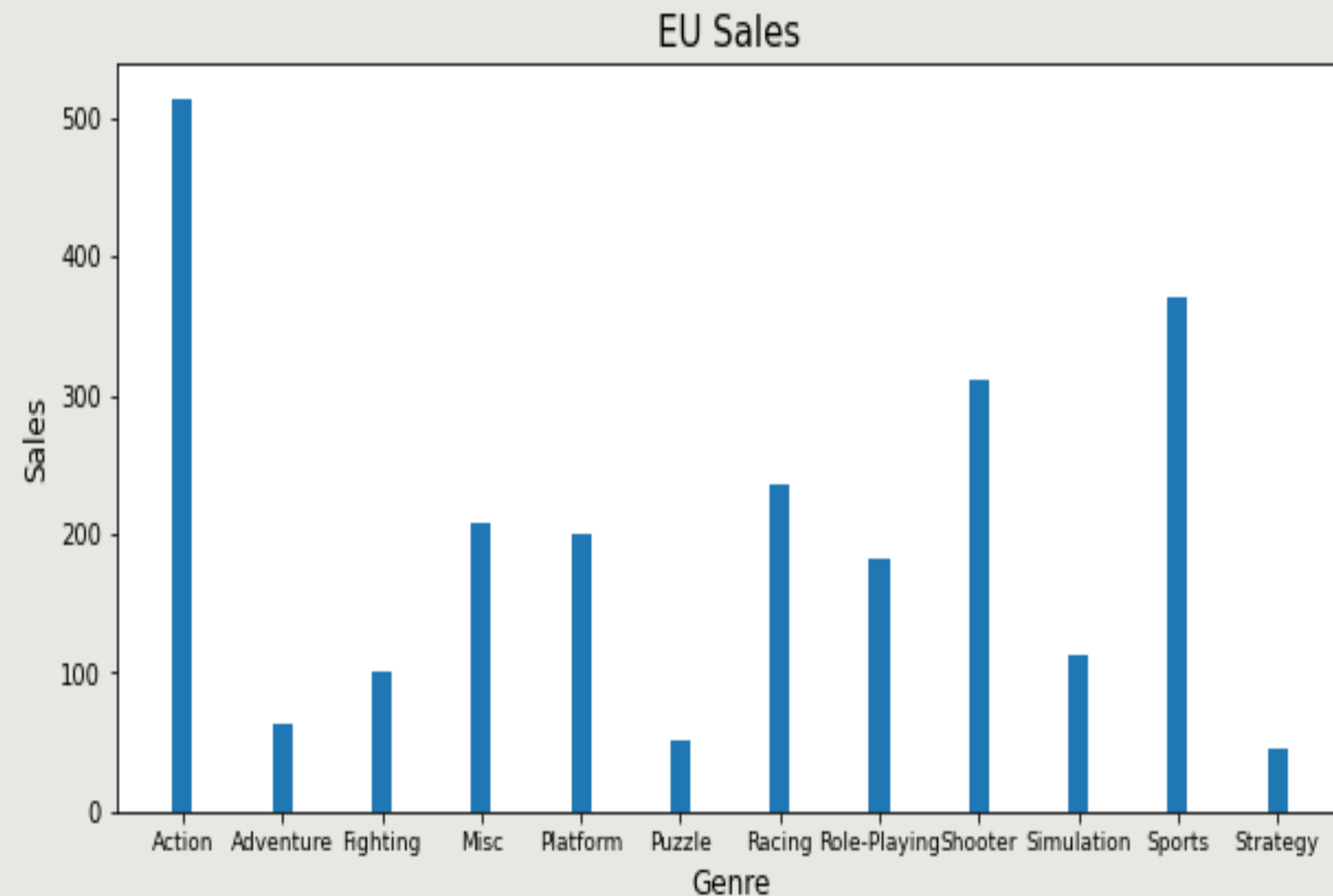
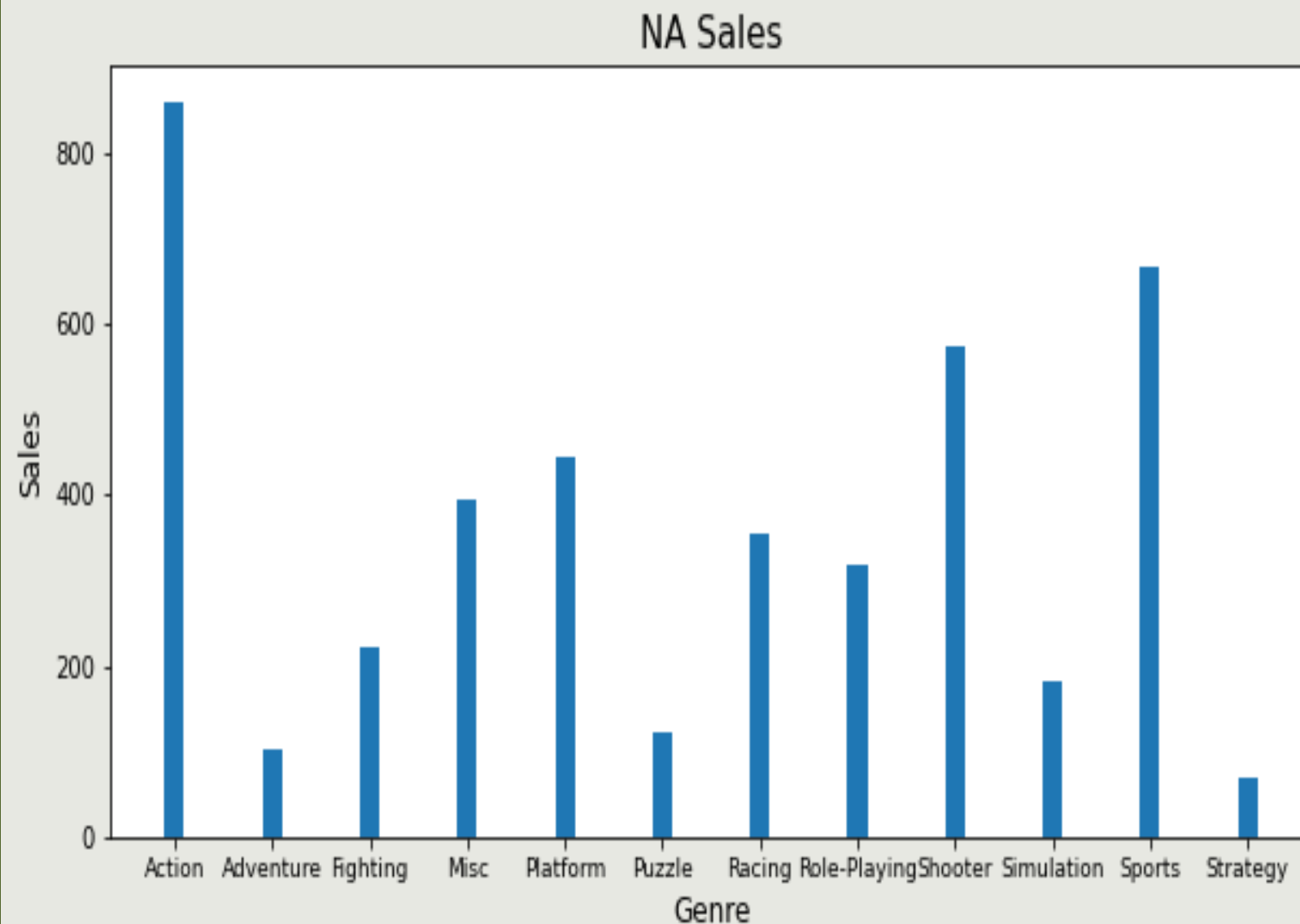
-> 각 지역별 출고량 데이터의 단위가 상이하야 단위를 million으로 통일

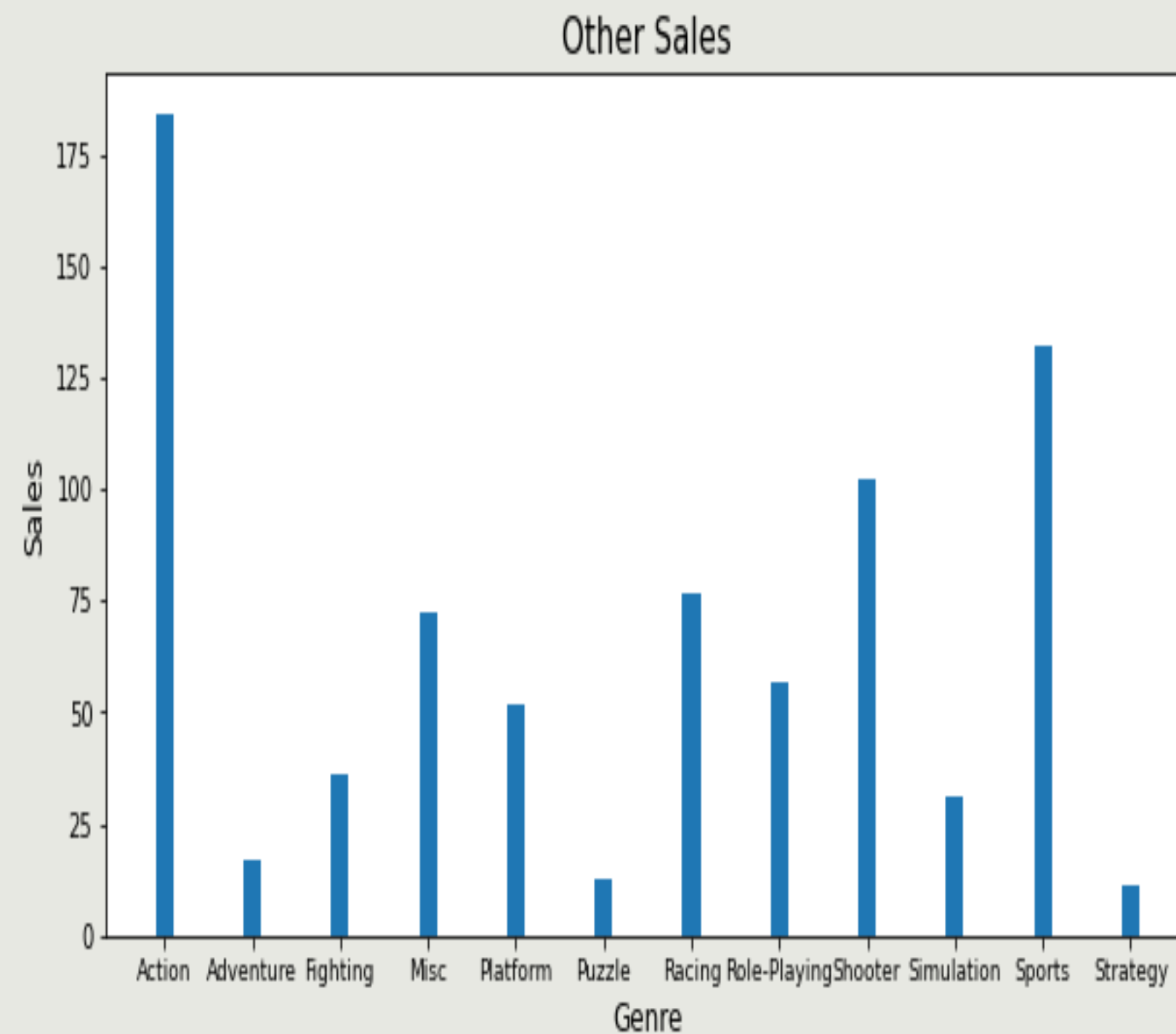
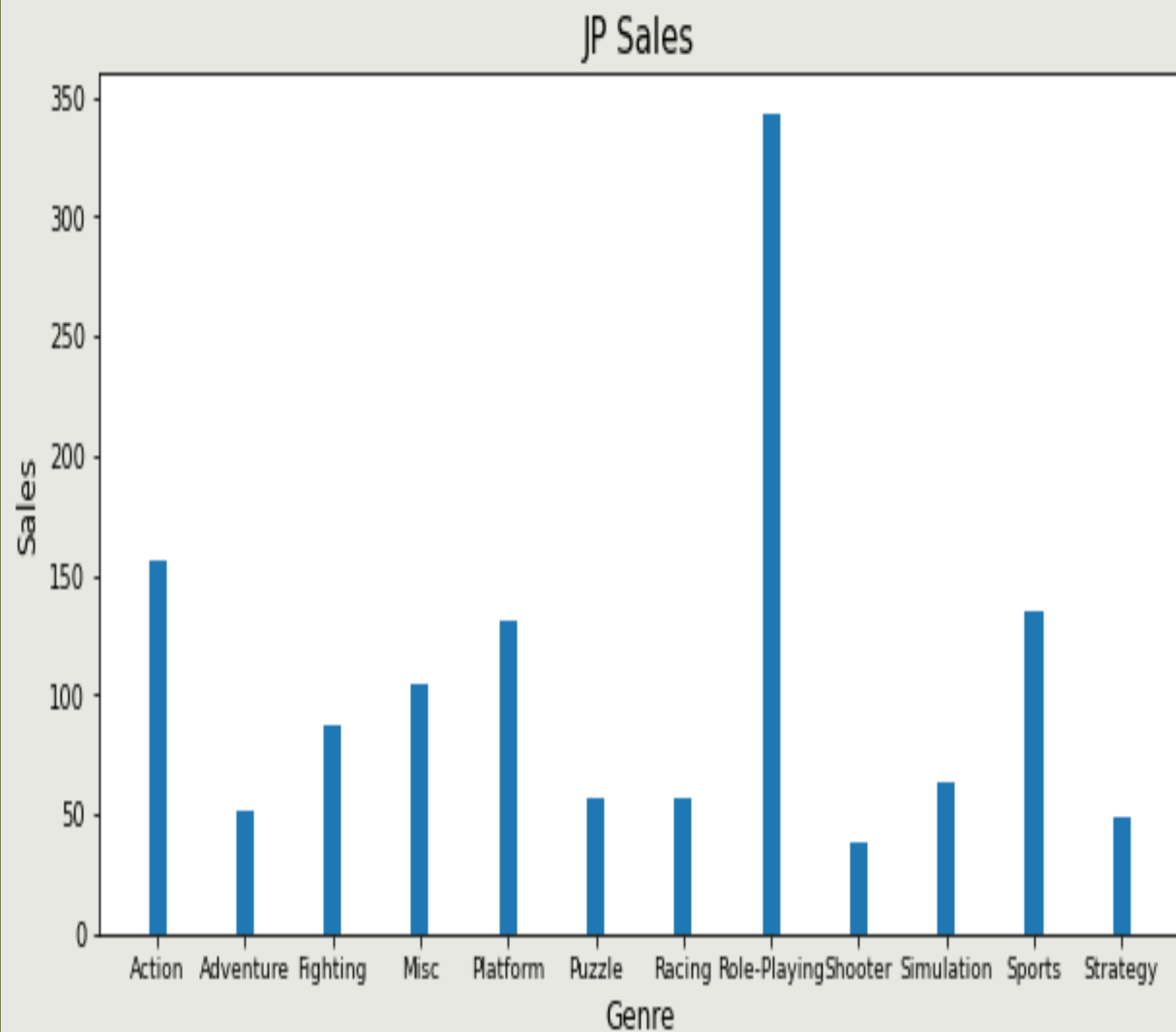
✓ 년도 표기 통일

-> 년도 데이터의 표기가 통일되어 있지 않아 천의 자리수까지 표기하도록 처리

✓ 지역별 출고량과 장르 두 변수 간의 연관성을 파악하기 위해 chi square test 검정 시행

▶ 결과: ‘장르와 지역별 출고량의 연관성이 있다’는 결론 도출





✓ bar plot을 통해 일본지역을 제외한 세 지역에서는 지역별 선호도가 유사한 양상을 띠고 있음을 파악: 1위 Action 2위 Sports 3위 Shooter

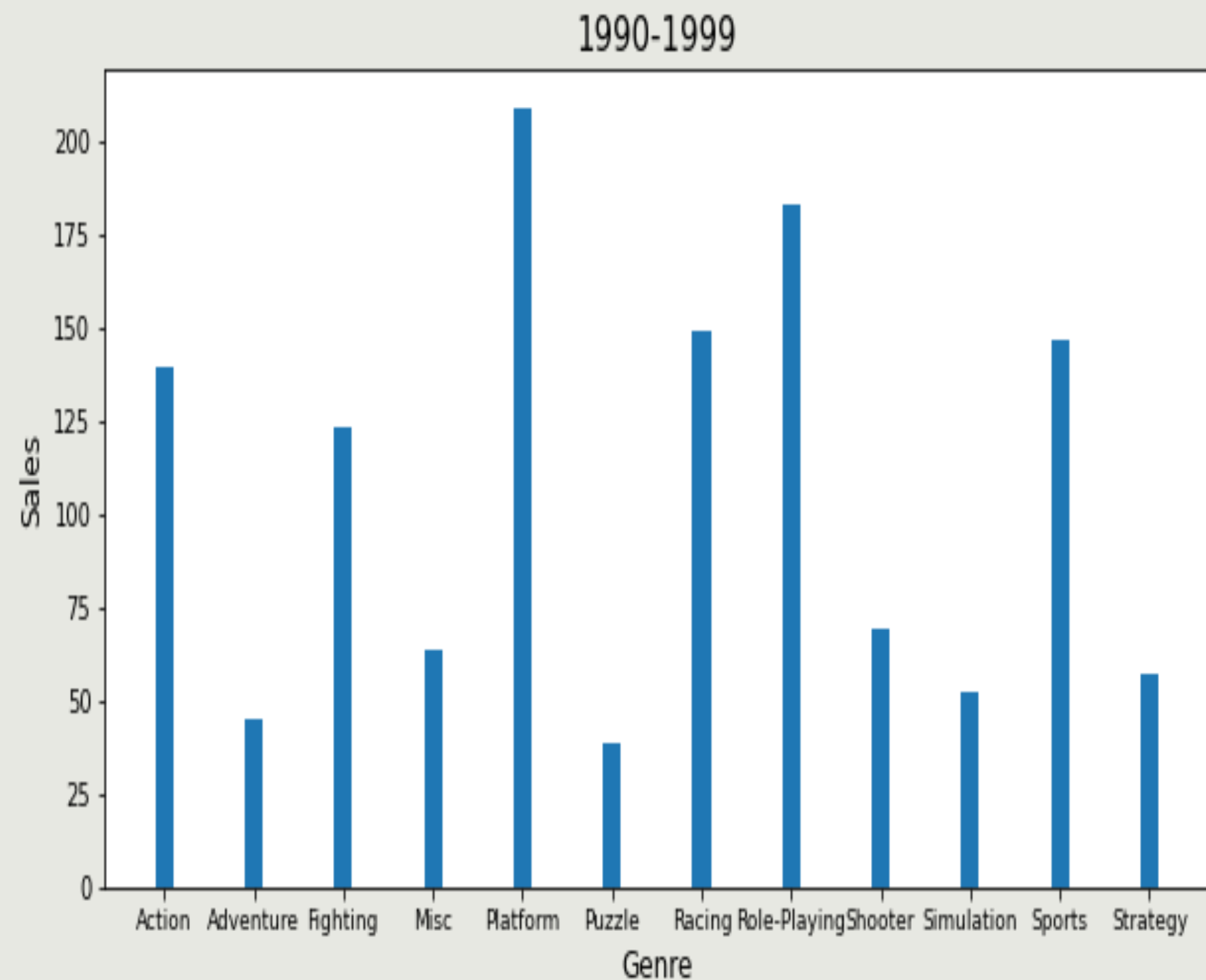
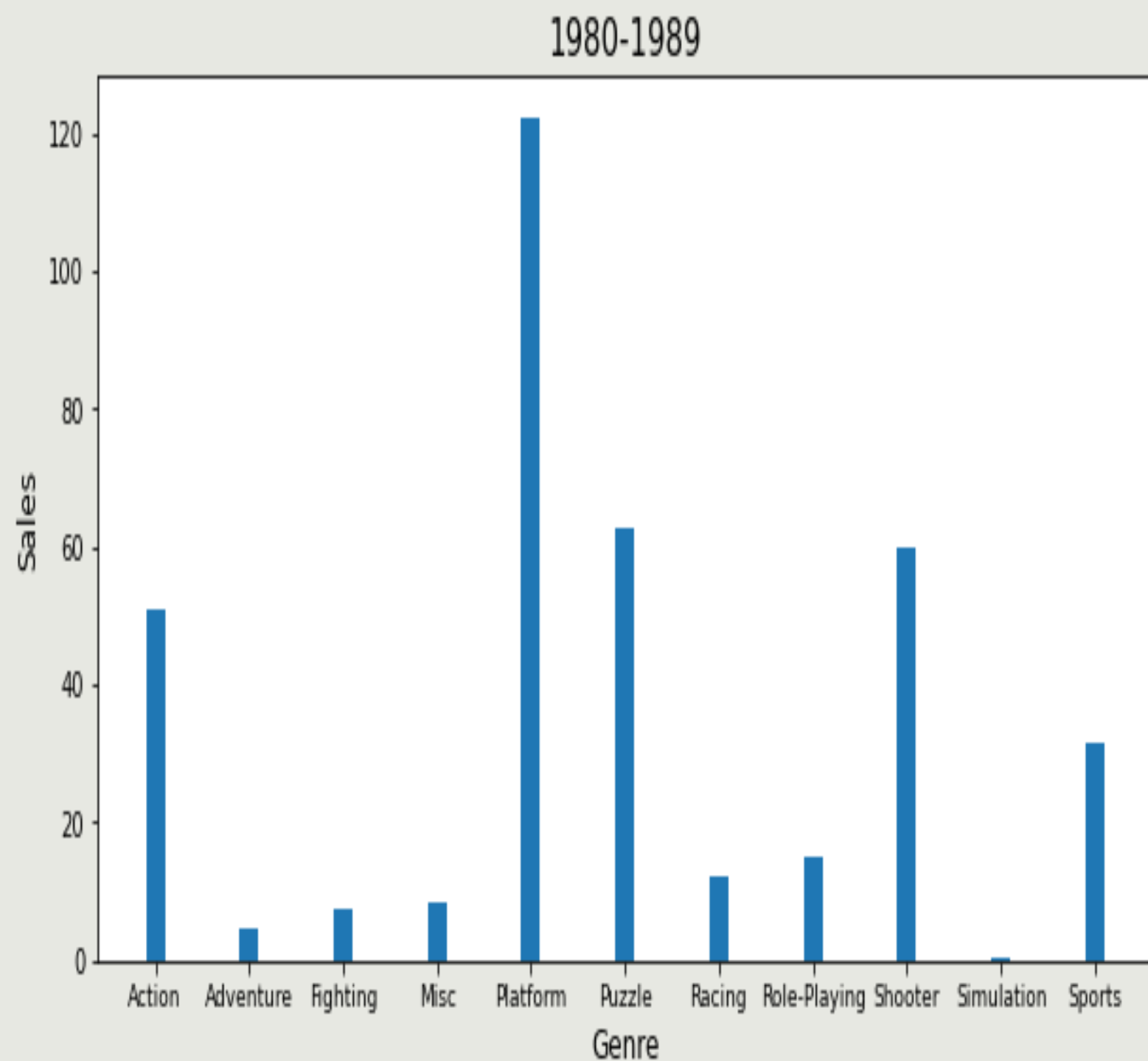
✓ 그러나 일본지역에서 세 지역과 장르 선호도에 차이가 있음:
1위 Role-Playing 2위 Action 3위 Sports

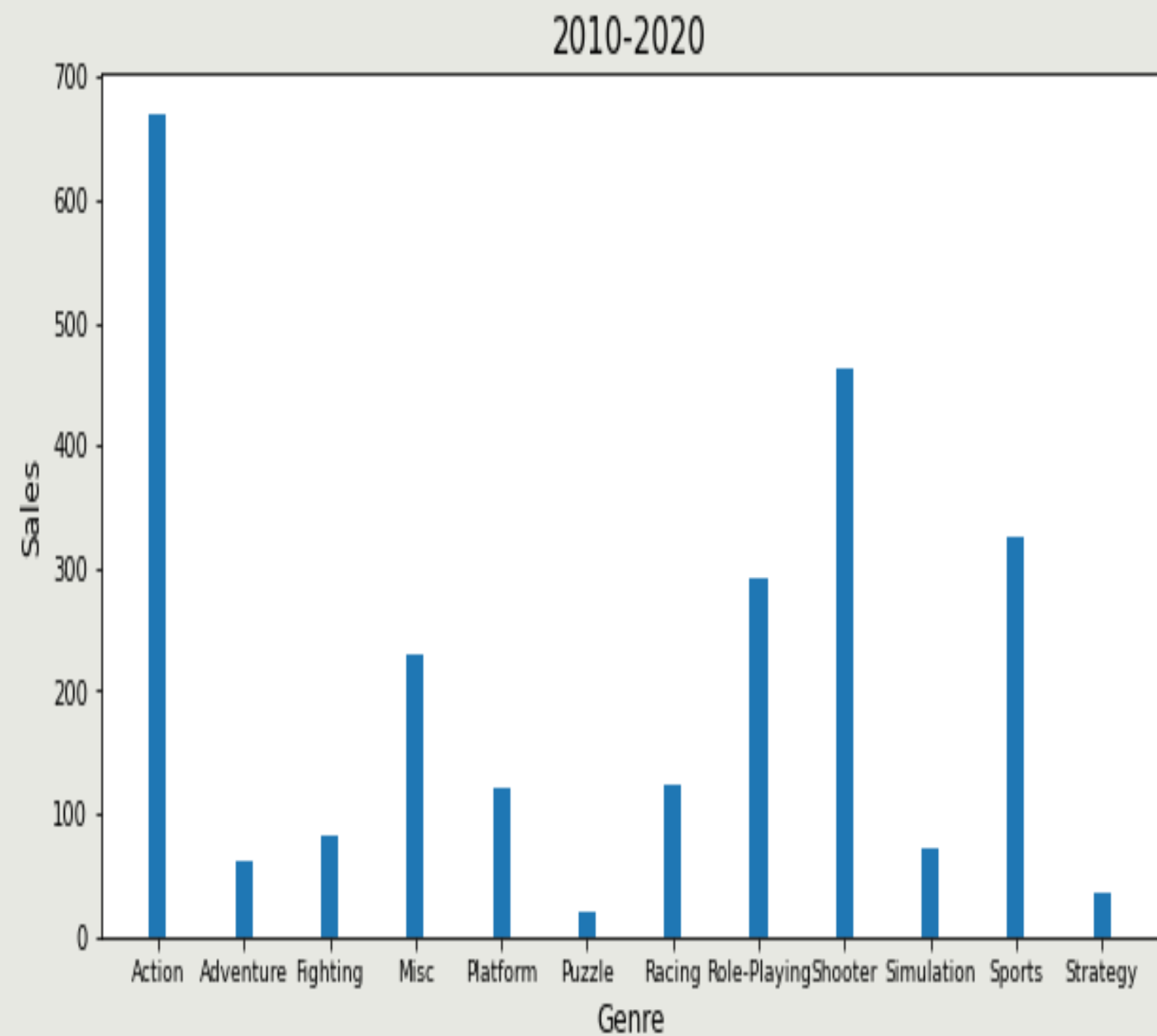
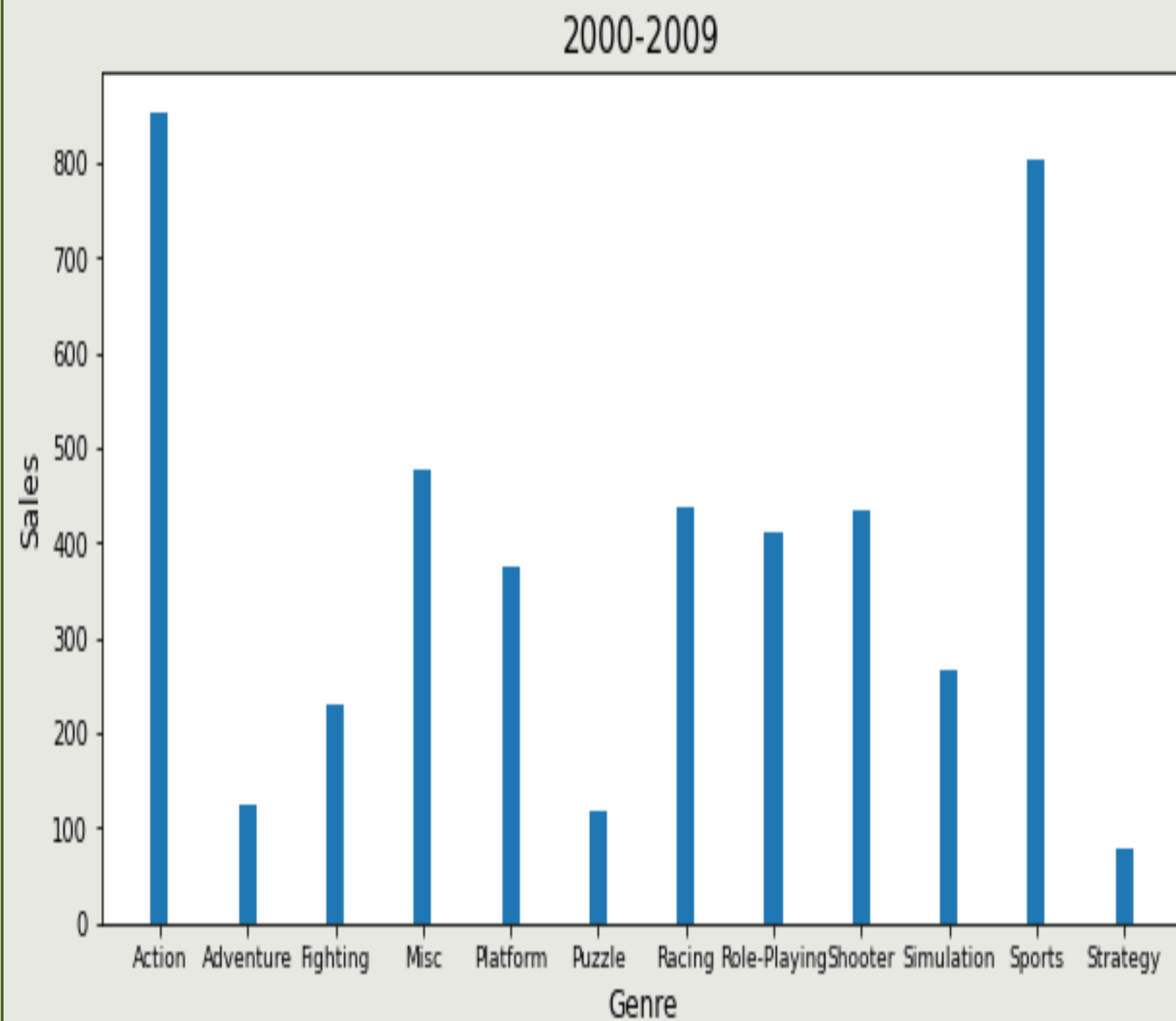
▶ **Role-Playing 게임(RPG)**: 참가자가 각자에게 할당된 캐릭터 (플레이어 캐릭터)를 조작하고, 일반적으로 서로 **협력하여 가상의 상황에서 주어지는 시련을 극복하면서 목표 달성을 목표로** 하는 게임의 일종

▶ 일본의 경쟁을 기피하고 혼자 즐길 수 있는 문화 생활을 즐기는 문화가 있어 Role-Playing 게임을 선호하게 되었다고 판단

▶ **결론: 지역별 게임 장르 선호도에 차이가 존재한다**

- ✓ 연도별 게임 트렌드가 있는지, 즉 연도와 장르 변수 간의 연관성을 파악하기 위해 chi square test 진행
- ✓ 결과에서 연도와 장르가 서로 연관되어 있음을 인지
- ✓ 1980~1989, 1990~1999, 2000~2009, 2010~2020별로 연도를 나누어 각 그룹에서 어떤 장르가 인기가 있는지 bar plot을 통해 진행
- ✓ 1980년대 strategy 장르의 출고량 데이터가 없어 결측치로 출력되나 0으로 대체





✓ 1980~1989: 1위 Platform 2위 Puzzle 3위 Shooter

✓ 1990~1999: 1위 Platform 2위 Role-Playing 3위 Racing

✓ 2000~2009: 1위 Action 2위 Sports 3위 Misc

✓ 2010~2020: 1위 Action 2위 Shooter 3위 Sports

▶ 20세기에는 platform 장르가 인기였고, 21세기에 접어들면서 action 장르로 관심이 쏠림

✓ 결론: '연도별 트렌드가 존재한다'는 결론이 도출

출고량이 높은 게임에 대한 분석

(1) Chi square test

✓ 총 출고량과 연도, 장르, 그리고 플랫폼과의 연관성을 살피기 위해 각각 세 가지 **chi square test** 시행

✓ 검정 결과

총 출고량
& 연도

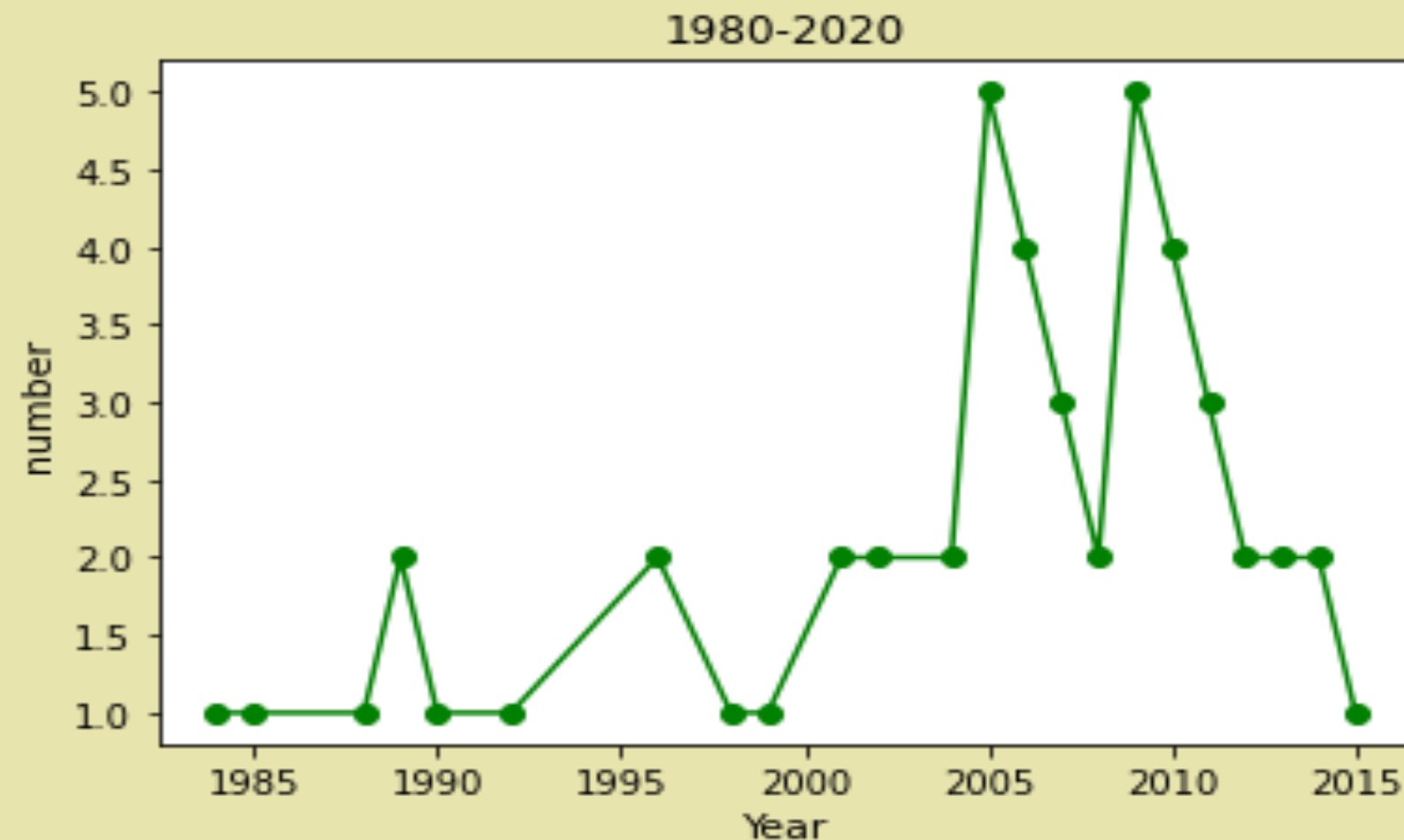
총 출고량
& 장르

총 출고량
& 플랫폼

▶ 연도, 장르, 플랫폼 모두 총 출고량과 연관성 있다

출고량이 높은 게임에 대한 분석 (2) line plot

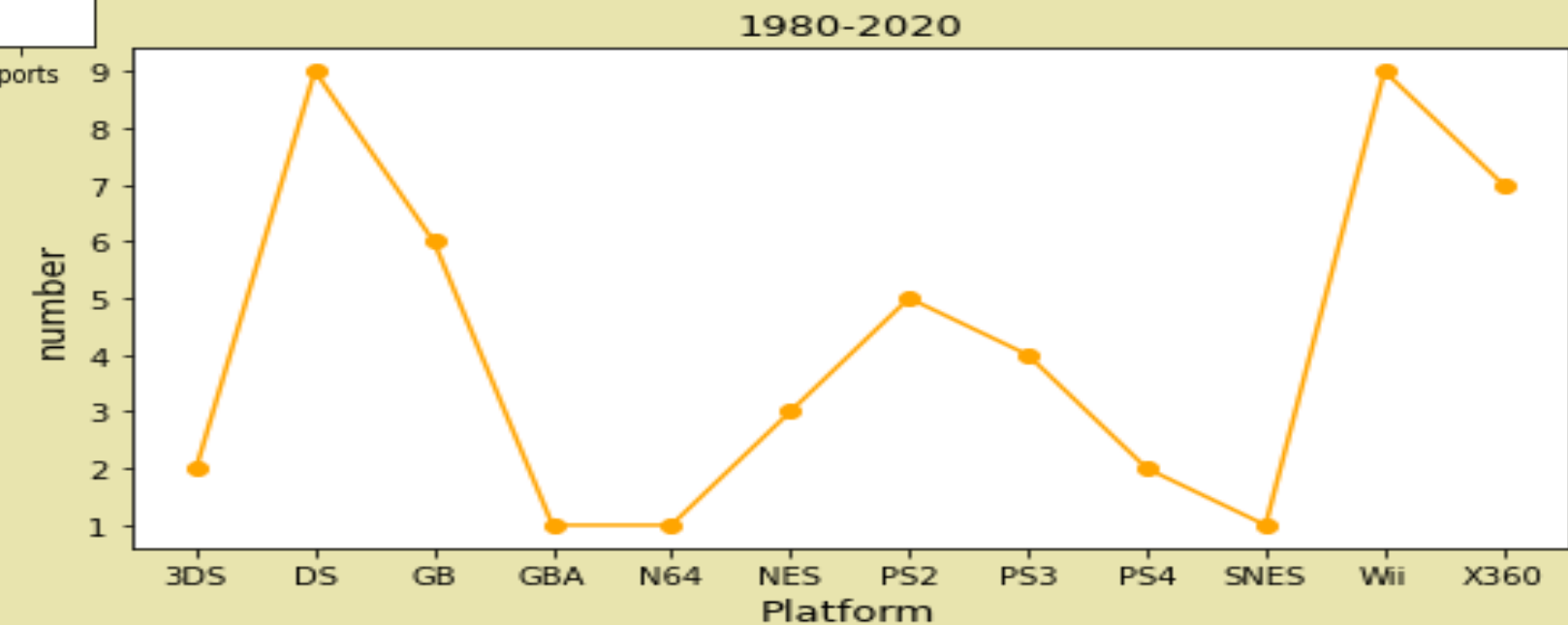
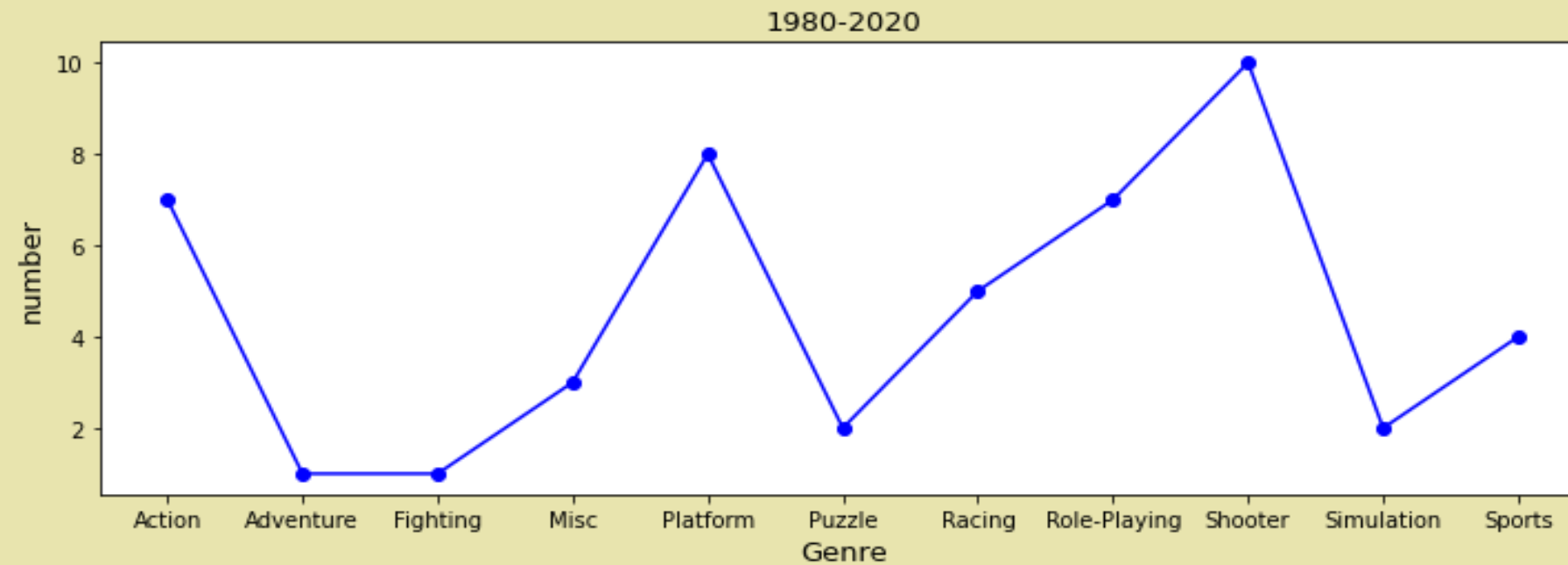
✓ 총 출고량 top 50 출시 년도를 살펴보면 2005~2010년까지의 게임들이 가장 많이 포함되어 있음



출고량이 높은 게임에 대한 분석 (2) line plot

✓ 장르를 살펴보면 **Shooter, Platform, Action** 장르 순으로 많음

✓ 플랫폼은 **Wii, DS, X360** 순으로 많음

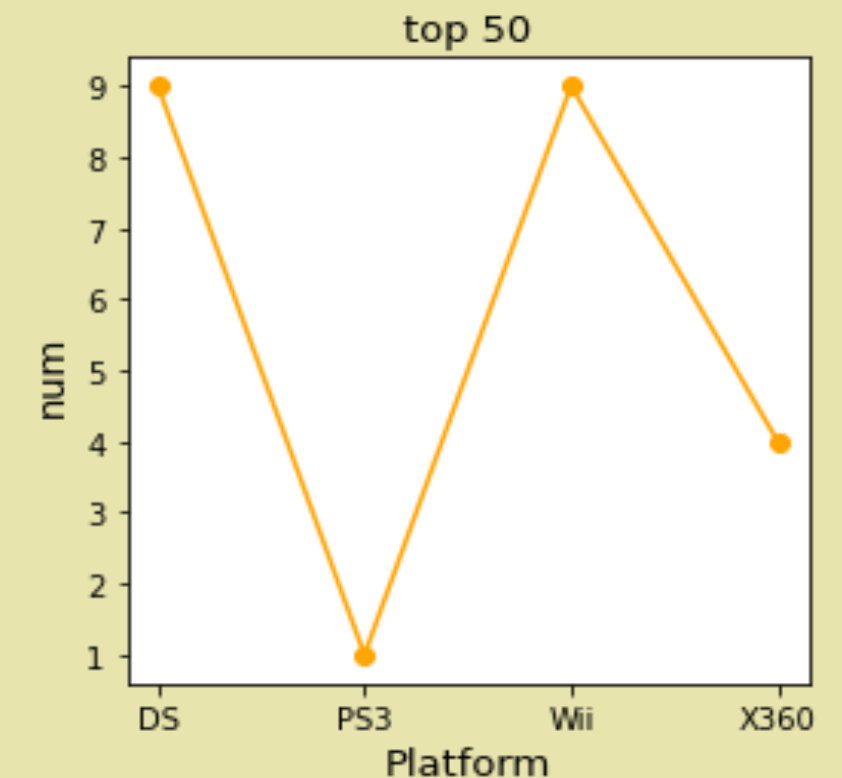
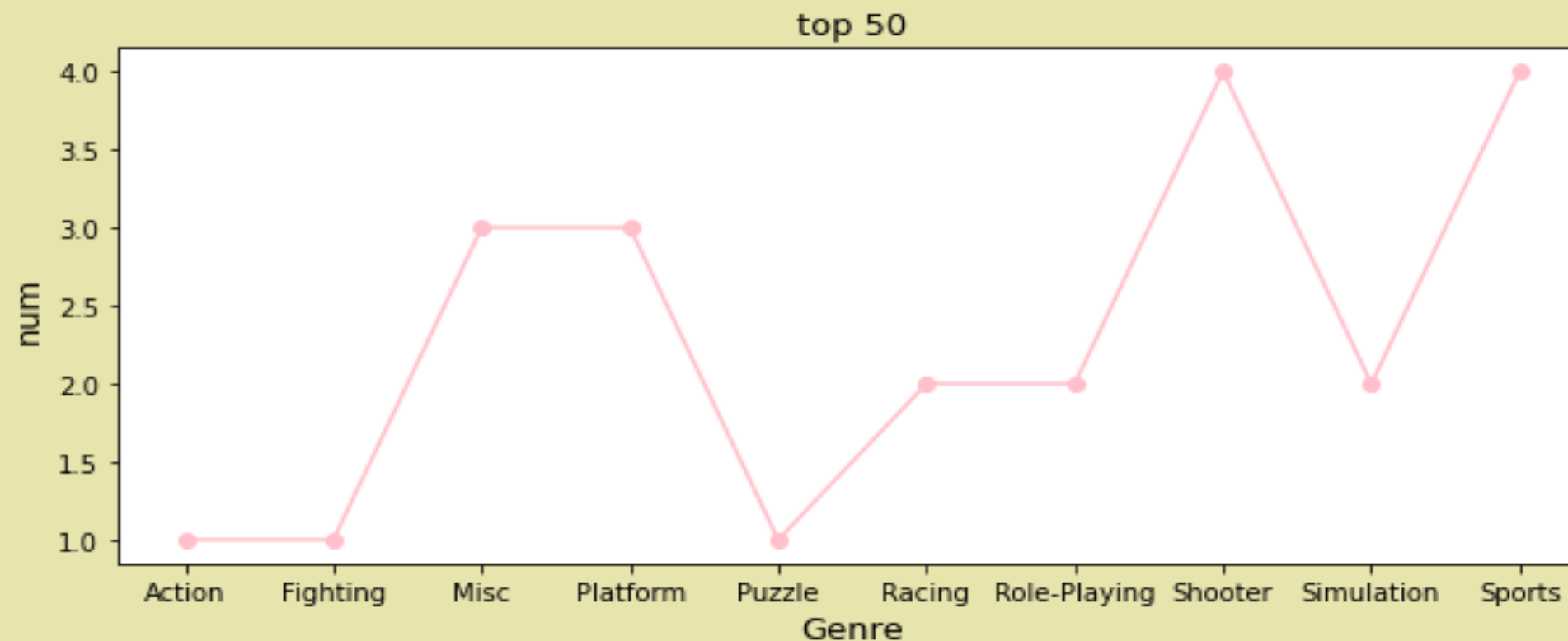


출고량이 높은 게임에 대한 분석 (3) 결론

✓ 1980~2020년에 출시된 게임 중 가장 많이 출고된 top 50개 중 다수가 2005년~2010년에 출시

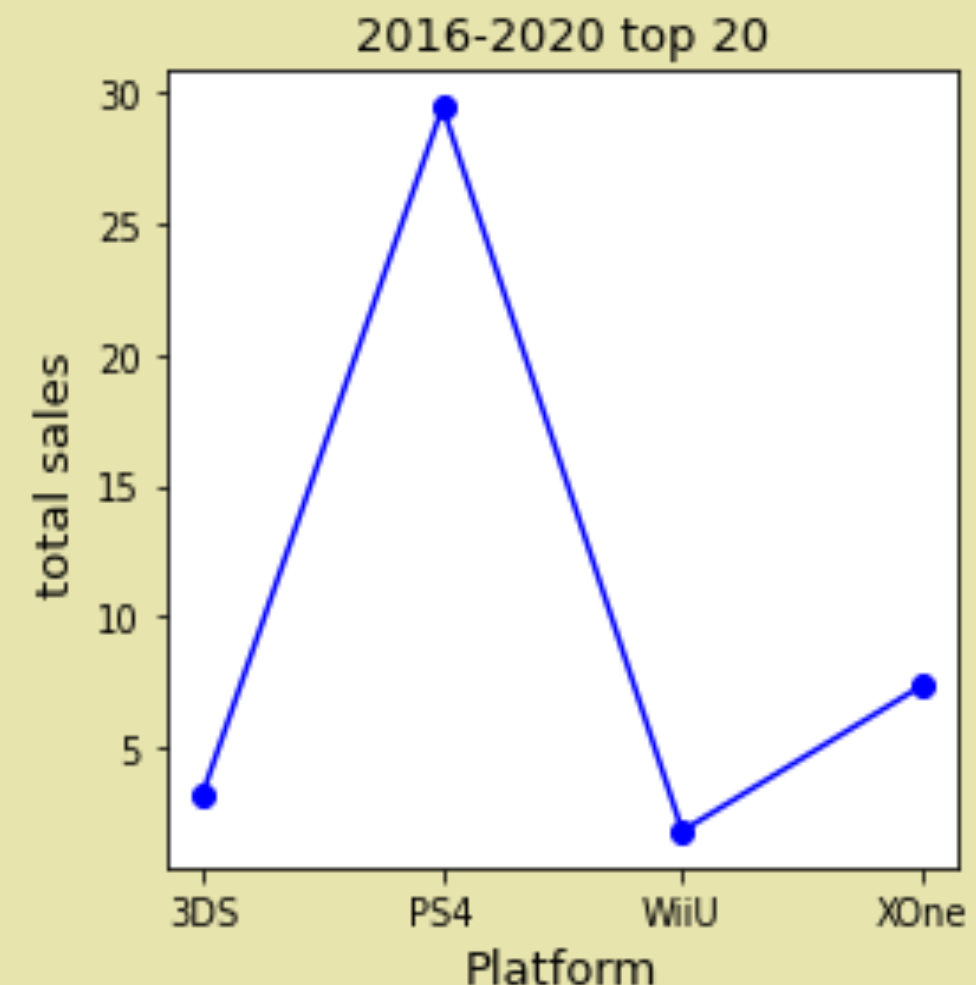
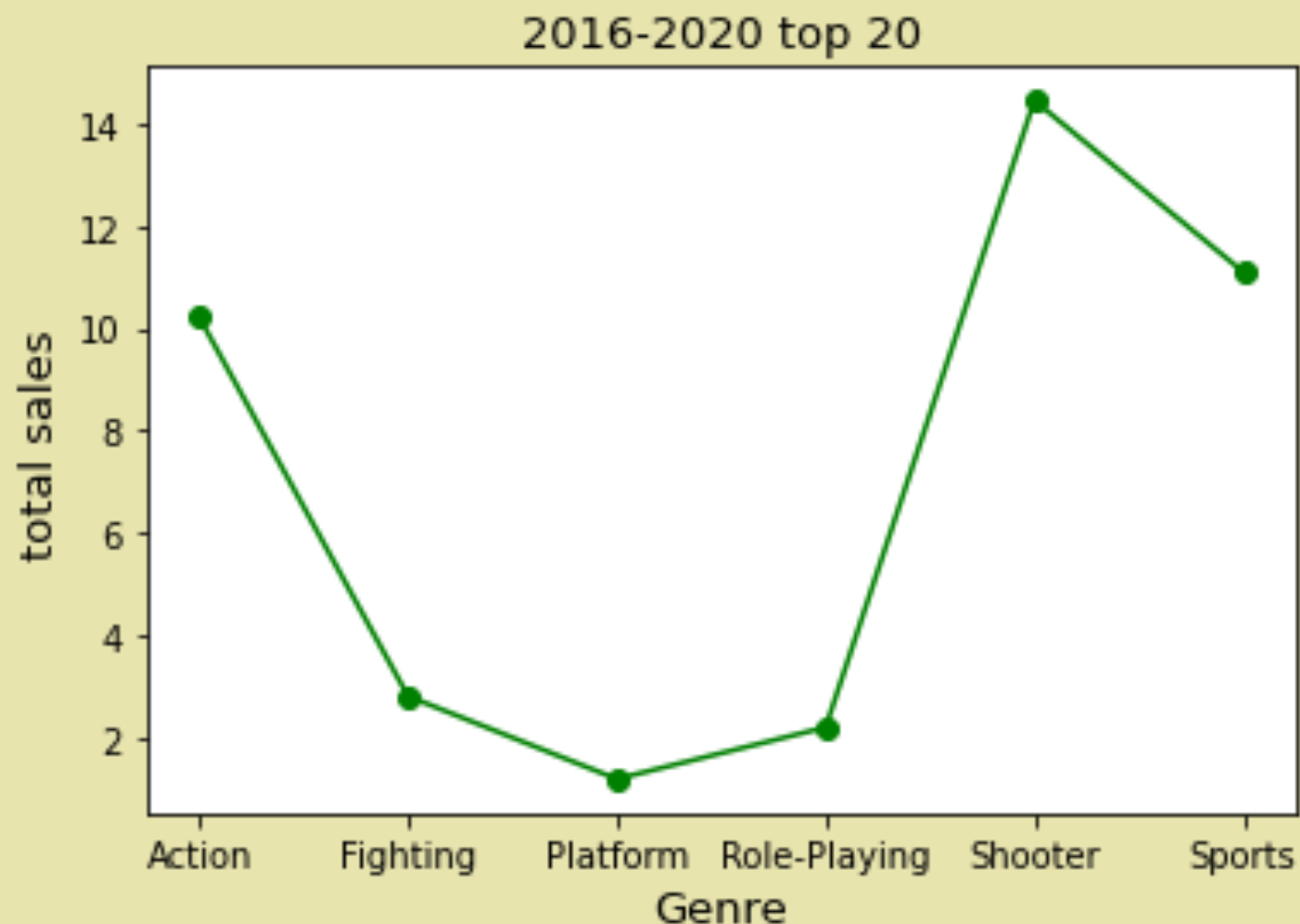
✓ 그 중 장르는 Shooter, Platform, Action 순으로 가장 많고, 플랫폼은 Wii, DS, X360 순

✓ top 50에서 2005~2010년 출시된 게임 중 다수는 Shooter나 Sports 장르이고, Wii 혹은 DS 플랫폼으로 출시



다음 분기에 설계해야 할 게임 (1) top 30의 장르&플랫폼 파악

- ✓ 최근 5년간 가장 많이 팔린 top 30에 대한 데이터 추출
- ✓ 총 출고량이 많은 장르는 Shooter, Sports, Action 순
- ✓ 플랫폼은 PS4, XOne 순으로 총 출고량 많음



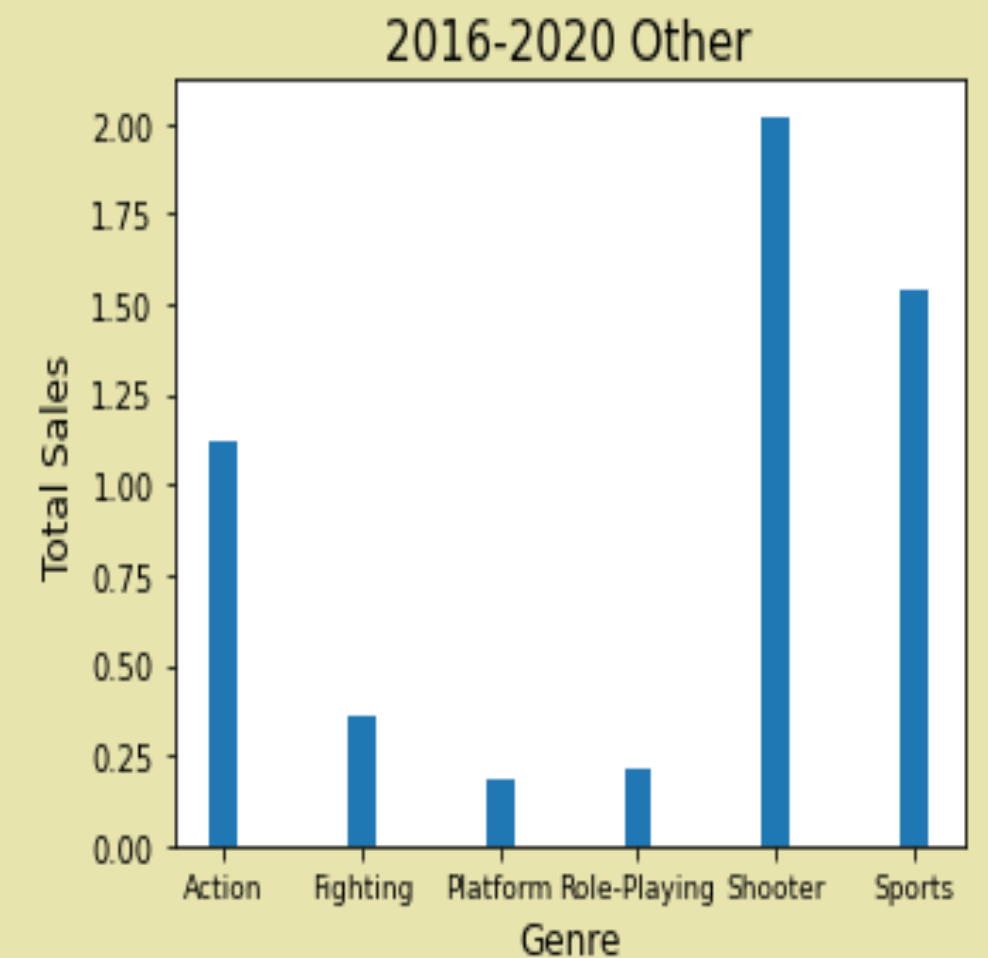
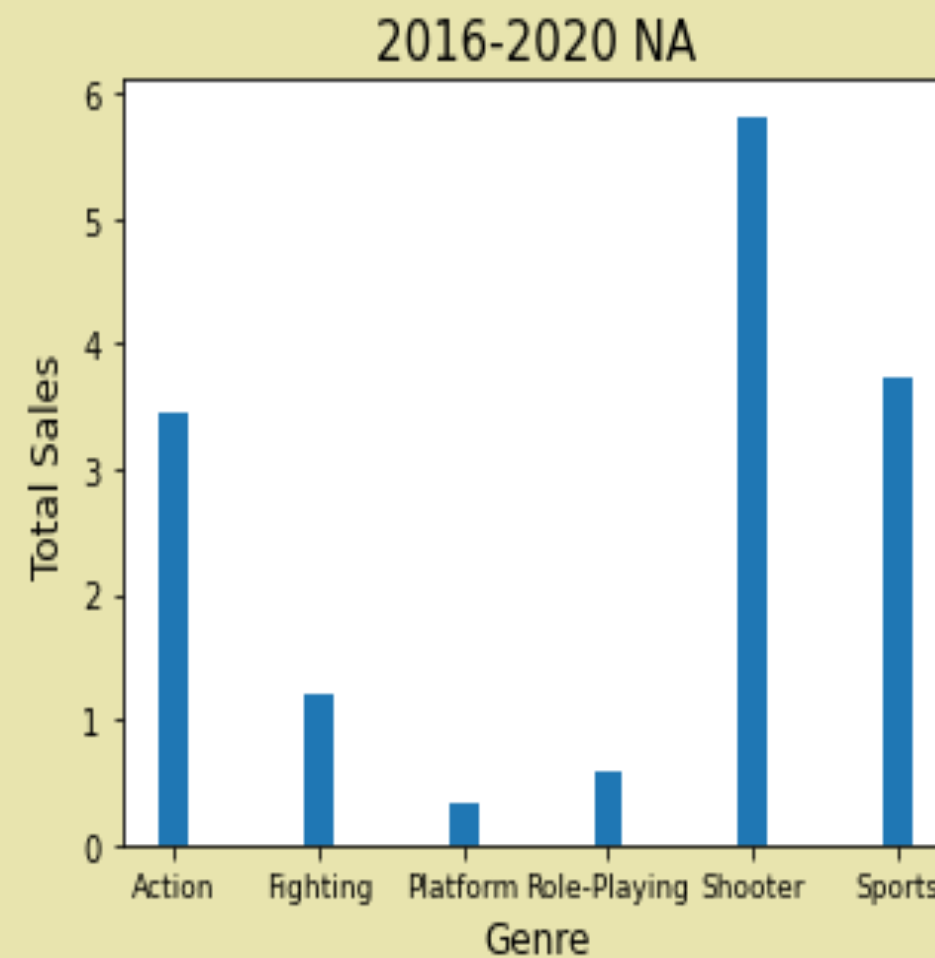
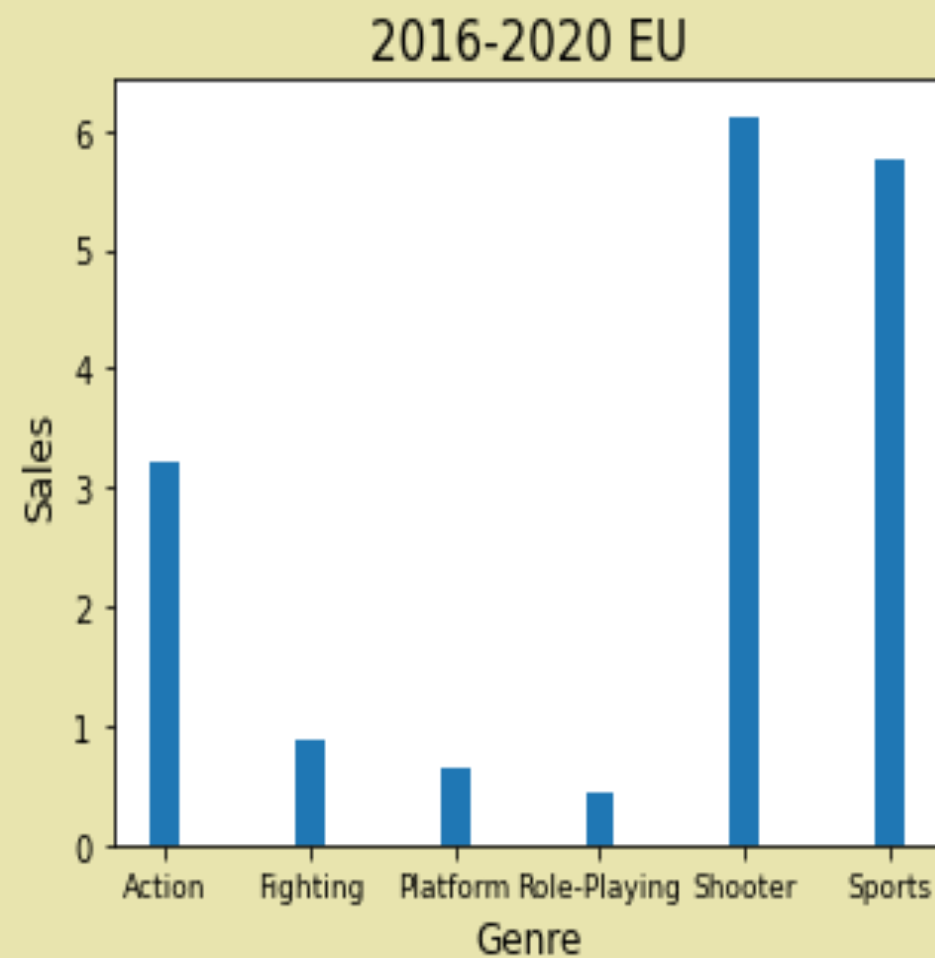
다음 분기에 설계해야 할 게임

(2) 각 지역 장르별/플랫폼별 출고량 연관성 확인

✓ 추출된 top 30 게임 데이터에 대해 각 지역별 출고량의 연관성을 확인하기 위해 chi square test 시행

✓ test 결과를 통해 기타 지역과 유럽 지역, 그리고 북미 지역과 유럽 지역 간의 출고량은 서로 연관되어 있다는 결론

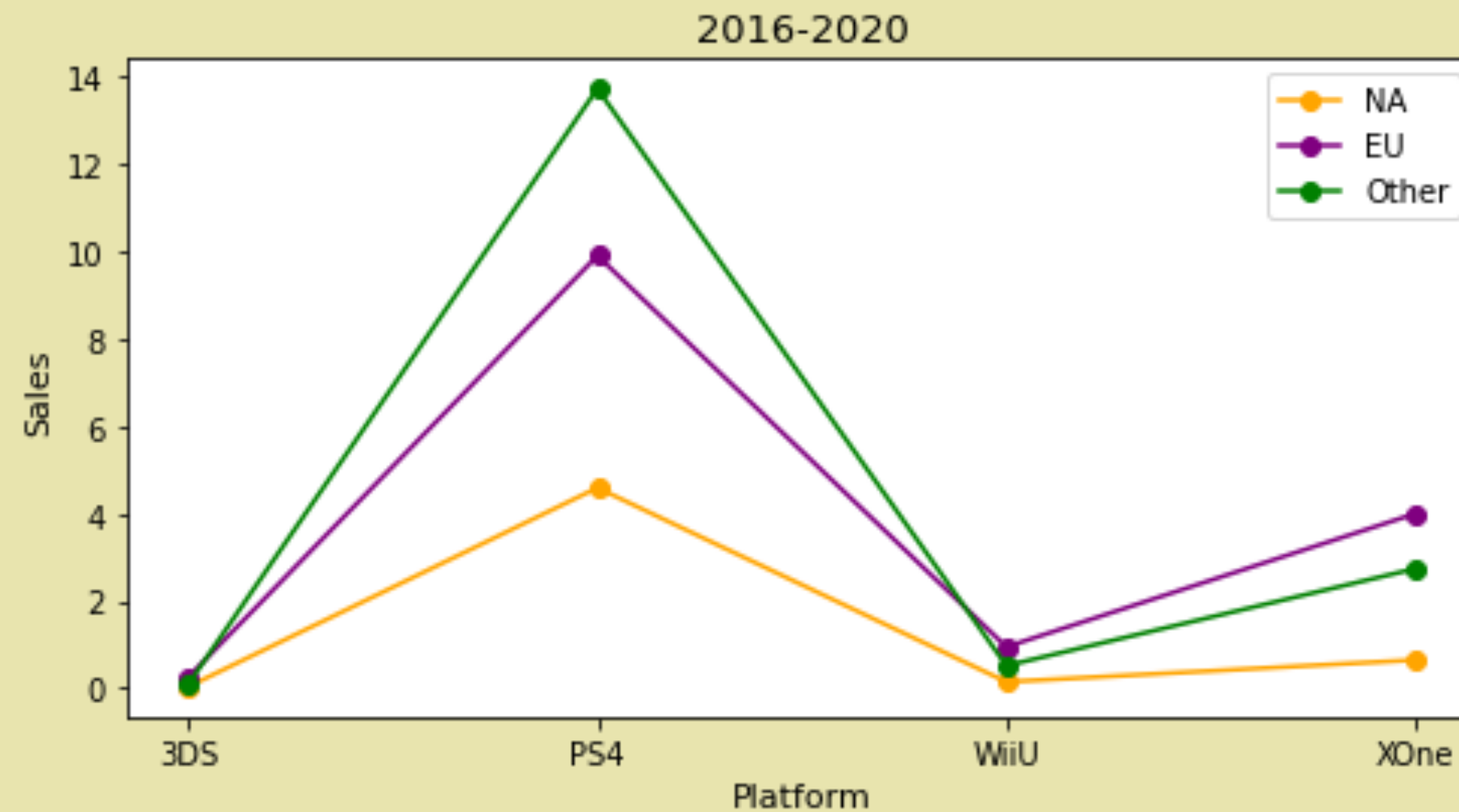
↓ bar plot을 통해 세 지역의 장르별 출고량이 유사한 형태를 보이고 있음을 알 수 있다



다음 분기에 설계해야 할 게임
(2) 각 지역 장르별/플랫폼별 출고량 연관성 확인

✓ 기타, 유럽, 북미 지역의 플랫폼별 총 출고량도 유사한 양상을 나타냄

↓ line plot



✓ 결론

▶ 비디오 게임 수요가 크게 감소하는 상황을 극복하기 위해서는 **대중성**을 공략하는 게임 필요

북미, 유럽, 기타 지역의 장르별/플랫폼별 선호도 양상이 유사

▶ 이 세 지역의 선호도를 만족하는 게임을 설계한다면 좋은 성과를 거둘 것이라 판단

▶ 세 지역의 선호도를 충족하기 위해 장르는 Shooter, Sports, Action 중 하나, 플랫폼은 PS의 최근 버전 PS5인 ‘대중적인 게임’을 설계해야 할 것

감사합니다.