**4주차 팀미션**

1. **슬랙 팀채널에서 팀원들과 함께 미션 토의를 진행해주세요.**
2. **리드부스터님께서 토의 후 내용을 취합하여 아래 템플릿에 맞게 팀 대표 미션을 작성해주세요.**
3. **슬랙 #00코치\_미션제출 채널에 팀 대표 미션을 제출해주세요.**

* **템플릿 확장자인 .docs로 제출해주세요.**

**Q1. 이번 미션에서는 데이터를 기본적으로 탐색하고, 간단한 통계와 시각화를 통해 데이터의 분포와 패턴을 파악하는 연습을 해보겠습니다. 수치형 데이터 분석, 카테고리별 트렌드 파악, 시간 기반 매출 분석 등을 단계별로 진행하게 됩니다. 이 과정을 통해 자연스럽게 데이터 분석 과정에 익숙해지고, 비즈니스 인사이트를 도출할 수 있는 능력을 키울 수 있을 것입니다.**

|  |
| --- |
| **[미션 진행 방법]**  **1. 데이터 다운로드:**  **[데이터 출처] \*다운로드 후 사용해주세요** [retail\_sales\_dataset.xlsx](https://docs.google.com/spreadsheets/d/1--NSL7U3MMABh_ILq44Xwezh6_Px5gIU/edit?usp=drive_link&ouid=101905656555390112166&rtpof=true&sd=true)  **2. Google Colab 복사본 생성**  - Colab이란? Google에서 제공하는 클라우드 기반의 무료 Python 개발 환경으로, 주로 데이터 분석, 머신러닝, 딥러닝과 같은 작업에 사용됩니다. Colab은 Jupyter Notebook의 형태를 따르며, 웹 브라우저에서 실행되기 때문에 복잡한 환경 설정 없이 바로 사용할 수 있다는 점에서 매우 편리합니다. Google Drive와 통합이 가능하며 공유와 협업이 가능하고, GPU와 TPU를 지원한다는 점이 가장 큰 장점입니다.  - 링크 접속: <https://colab.research.google.com/drive/1PmrAHwlYpAEB8rs2h8I8LRGVJ7A07Bs2?usp=sharing>  - 우측 상단 ‘파일’ > ‘Drive에 사본 저장’을 클릭해주세요. (개인 Drive에 파일 복사본 생성)    **3. Google Colab 사용법**    1. “+ 코드”: 코드 셀 추가  2. “+ 텍스트”: 마크다운(markdown) 셀 추가  3. ➲: 코드 실행버튼 (cmd+enter 혹은 ctrl+enter로도 실행 가능합니다.)  4. “셀”: 하나의 회색 박스를 “셀”이라는 단위로 부릅니다.  5. “#주석”: 코드를 설명하기 위한 수단으로 ‘#’로 시작하는 주석을 작성할 수 있습니다.  **6. 데이터 불러오기: pd.read\_excel( ), pd.read\_csv( ) 등의 함수를 사용할 수 있습니다.**  **경로를 설정할 때에는 pd.read\_excel(‘폴더/파일.xlsx’)처럼 작성해줍니다.**  7. 과제를 수행하실 때에는 해당 박스에 코드를 채워 넣으신 후, 실행을 해서 제출해주시면 됩니다.  **➡ 1~3번 가이드 영상**  [4주차 미션\_경로 복사.mp4](https://drive.google.com/file/d/1CqBdirLHIvdswUG-VLmbSjEO-qPayXLA/view?usp=drive_link) |

|  |
| --- |
| **[미션]**   1. 데이터 기본 탐색 및 EDA(Exploratory Data Analysis)    * 데이터 구조 및 분포      + 데이터셋의 행과 열의 수를 출력해주세요.      + "Quantity" (수량), "Price per Unit" (단가), "Total Amount" (총 금액) 같은 수치형 컬럼에 대한 기본 통계(평균, 중앙값, 최소값, 최대값)를 구해보세요.      + 몇 명의 고유 고객이 존재하나요? 고객의 연령 분포는 어떻게 되나요? 평균 연령과 표준 편차를 계산해주세요.    * 카테고리와 트렌드 탐색      + 몇 개의 고유한 제품 카테고리(Product Category)가 존재하며, 각 카테고리의 빈도는 어떻게 되나요? 총 매출 기준으로 가장 많은 판매가 이루어진 카테고리는 무엇인가요?      + 카테고리별 "Quantity" (수량)의 평균은 어떻게 되나요? 카테고리별로 구매량에 패턴이 있나요?      + 성별에 따른 카테고리별 구매 패턴은 어떻게 되나요? 특정 제품이 특정 성별에 더 많이 판매되나요?    * 시간 기반 분석      + 가장 빈번한 거래 날짜는 언제인가요? 특정 월에 매출 피크가 나타나나요?      + 월별 또는 계절별 고객 구매 패턴에서 어떤 트렌드를 발견할 수 있나요? "Total Amount" (총 금액)가 월별로 큰 차이를 보이나요? 2. 심화 분석 및 시각화    * 매출 트렌드 분석 및 예측      + 전체 매출의 몇 퍼센트가 각 제품 카테고리에서 발생하나요? 어떤 카테고리가 가장 많은 매출을 기여하고 있나요?      + 데이터에서 발견되는 트렌드를 바탕으로 미래 고객 행동이나 판매 패턴에 대한 예측을 할 수 있나요? 매출에 가장 큰 영향을 미치는 요인이 무엇인지 도출할 수 있나요? (예: 제품 카테고리, 고객 연령, 거래 날짜 등)    * 핵심 인사이트 시각화      + 제품 카테고리별 매출 비율을 보여주는 파이 차트 또는 바 차트를 만들어보세요.      + 고객의 연령 분포를 히스토그램으로 시각화해보세요. 특정 연령대가 주로 구매를 하는지 확인할 수 있나요?      + 총 매출이 시간에 따라 어떻게 변화하는지 라인 차트로 시각화해보세요. 월별로 매출 트렌드가 명확하게 보이나요?      + 고객 연령과 총 지출 사이의 관계를 나타내는 산점도를 만들어보세요. 고가 고객의 그룹을 식별할 수 있나요? 3. 추가 분석    * 혹시 추가로 진행한 분석이 있다면, 함께 작성해주세요. |

|  |
| --- |
| **[제출 양식]**   * <https://colab.research.google.com/drive/1PmrAHwlYpAEB8rs2h8I8LRGVJ7A07Bs2?usp=sharing> * **반드시 제출 양식을 복사한 이후 작업 진행해주세요.**   **파일 > Drive에 사본 저장** |

|  |
| --- |
| **[코랩 링크 제출]**  : https://colab.research.google.com/drive/1R4GI2CMth0wSCSPXs5\_lpYq6BW0ncWob?usp=sharing |

|  |
| --- |
| [평가 기준]   * 결과값이 정확한지 평가. * 데이터 분석/시각화를 기반으로 분석 인사이트가 적합한지를 평가. * 핵심 인사이트 시각화 파트에서 적절한 그래프로 시각화 결과가 출력되었는지를 평가. |

**Q2. 다양한 태스크에 특화된 GPT 모델을 찾아볼 수 있는 플랫폼, GPTs 스토어에 대해 알아보았습니다. 각 GPTs는 특정 작업이나 주제에 최적화되어 있고, 이를 통해 사용자는 자신의 필요에 맞는 GPT를 선택해서 업무 생산성, 창의적 작업, 데이터 분석 등 다양한 분야에서 활용할 수 있다고 합니다. GPT 스토어에 익숙해질 수 있도록 함께 실습을 진행해보겠습니다.**

|  |
| --- |
| 1. 자신에게 필요한 GPT를 GPT 스토어에서 선택하세요. 2. 선택한 GPT에 대해 간단히 소개하세요. 3. 데이터 분석/시각화를 위한 유스 케이스를 제작하고, 이를 위한 프롬프트 예시 3가지를 제시하세요. |

예시)

|  |  |
| --- | --- |
| GPTs | 회의록 작성 GPT |
| 유스케이스 | 회의 기록 및 요약 자동화. 회의 중 자동으로 핵심 내용을 요약 및 정리하여 팀에 배포. |
| 프롬프트 예시 | "지난 1시간 동안의 회의 대화를 요약하고, 결정된 주요 사항을 나열해 주세요." |
| "다음 회의 전까지 해결해야 할 주요 액션 아이템을 목록으로 정리해 주세요.” |
| "회의 중 나온 의견과 그에 따른 결론을 간단히 요약해 주세요." |

[제출물]

|  |  |
| --- | --- |
| GPTs | 고객 행동 분석 GPT |
| 유스케이스 | 고객의 구매 패턴 분석 및 예측. 고객의 행동 데이터를 바탕으로 구매 트렌드를 파악하고, 향후 행동을 예측하여 마케팅 전략에 활용. |
| 프롬프트 예시 | **고객의 구매 연령대 분석** "고객 데이터를 바탕으로, 연령대별 구매 패턴을 분석하고, 각 연령대의 총 구매액을 비교해 주세요. 또한, 이를 바탕으로 연령대별 주요 구매 제품을 시각화해 주세요." |
| "고객의 구매 이력을 바탕으로, 향후 3개월 동안 가장 많이 팔릴 제품 카테고리를 예측해 주세요." |
| "과거 3년간의 월별 매출 데이터를 기반으로 계절별 매출 변동 패턴을 분석하고, 향후 6개월 간의 매출 예측을 해 주세요. 예측 결과를 선 그래프로 시각화해 주세요." |

|  |
| --- |
| [평가 기준]   * 데이터 분석, 시각화에 특화된 GPTs를 선택했는지를 평가. * 데이터 분석, 시각화에 맞는 유스케이스를 제안했는지를 평가. * 위에서 답변한 유스케이스에 맞는 프롬프트를 3가지 작성했는지를 평가. |