Track 1. 부문서류

|  |  |
| --- | --- |
| Track 1.  ChatGPT 활용 사례 부문  ※ 본 파일은 참가자에게 공개되는 양식 원본입니다.  ※ 본 파일을 다운 받거나 사본으로 복사한 뒤 작성해주세요. | |
| 지원 기간 | 2024년 11월 20일 00:00 ~ 2024년 12월 15일 23:59 (한국 시간 기준)까지 |
| 지원 방법 | 이메일 접수 ( gala2024@day1company.co.kr ) |
| 필수 제출 서류 | [1] 공통 제출 서류  [2] 부문별 제출 서류 ◀ 현재 보고 계신 본 서류입니다.  [3]부문별 별도 제출 항목 \* 지원자들은 공모작이 제3자의 소유권, 저작권, 초상권 등 기타 제반 권리를 침해한 표절, 도용, 모방작이 아님을 보증해야 합니다.  \* 공모작에 사용되는 모든 폰트 및 에셋, 음원 등 공모작에 활용한 요소들은 모두 상업적 이용이 가능한 자료를 사용해야 합니다. (필요 시 저작권자에게 이용 허락을 사전에 지원자가 직접 구해야 합니다.)  \* 위 사항을 포함, 관련하여 발생하는 모든 분쟁에 대한 법적 책임은 모두 지원자에게 있으며, 지원자의 귀책사유로 인하여 패스트캠퍼스(주식회사 데이원컴퍼니)와 제3자 간에 분쟁이 발생한 경우 지원자는 자신의 책임과 비용으로 패스트캠퍼스를 적극 면책시키고, 컴퍼니가 입은 손해(변호사 보수를 포함하되 이에 한하지 않음)를 배상하여야 합니다. |
| **[본 파일 해당]**  파일명 규정 | **지원분야\_개인/팀\_참가팀명\_부문서류**  ex. Track1\_팀\_패캠GALA팀\_부문서류  ex. Track1\_개인\_패캐머\_부문서류 |
| **[본 파일 해당]**  **필수 작성 항목** (세부 내용 아래 참고) | **1. 참가명/참가팀명**  **2. 공모작 명**  **3. 공모작 한 줄 소개글**  **4. 활용 배경/목적**  **5. 사용 프롬프트**  **6. 시행착오와 해결 방안**  **7. 생산성 향상 정도**  \* 서식의 변경 및 삭제는 불가하며 각 항목별 지정된 가이드를 엄수해주세요.  \* 작성한 본문을 드래그 한 뒤 도구 - 단어 수 (Ctrl + Shift + C)를 통해 현재 글자 수를 확인할 수 있습니다. |
| **[본 파일 해당]**  선택 작성 항목 | 1. 사용한 실제 프롬프트 링크 : 제출 시 심사에 더 도움이 됩니다 (필수 4번 연계)  2. 나만의 활용 노하우 (공모작과 관련한 재미있는 에피소드가 있다면 작성)  3. 추가 사용한 보조 AI 프로그램 (ChatGPT 외에 보조적으로 사용한 도구가 있다면 작성) |

## I. 공모작 요약 [필수작성]

\* 참고사항 : 공모작 요약 내용은 본선 진출 시, 투표를 위한 자료로 사용 될 예정입니다.

데이터 분석 과제를 ChatGPT로 최적화하여 생산성과 완성도를 극대화

|  |  |
| --- | --- |
| **1. 참가명 / 참가팀명** | { 공백 포함 20자 이내로 작성 }  - 개인 참여 시에도 참가명을 필수로 입력해주세요.  - 동일 분야 동일 인물의 중복 참여 시 참가명/참가팀명을 다르게 입력해주세요. |
| 천영진 | |
| **2. 공모작 명** | { 공백 포함 30자 이내로 작성 }  콘텐츠를 대표할 수 있는 제목을 작성해주세요. |
| 빅데이터과 1학년에게 GPT의 의미 | |
| **3. 공모작 한 줄 소개글** | { 공백 포함 50자 이내로 작성 }  결과물과 관련한 본인 만의 WoW Moment를 중심으로 작성해주세요. |
| ChatGPT로 방대한 데이터 분석 효율적 창의적 인사이트 도출 | |

## II. 공모작 세부 [필수작성]

|  |  |
| --- | --- |
| **4. 활용 배경/목적** | { 공백 포함 300자 이상 작성 }  활용 배경/계기 또는 목적, 사용방식에 대해 최대한 자세하게 작성해주세요. |
| 저는 대학교 1학년으로서 아직 Python을 거의 구사하지 못하는 상황이었고, 과제를 내주신 교수님도 이러한 점을 인지하고 있었습니다. 또한 교수님은 평소에 생성형 ai의 적극 활용을 매우 강조하셨던 분이셨고 따라서 교수님께서는 항상 생성형 AI를 활용해야만 완성을 겨우 할 수 있을 정도로 난이도가 높은 과제를 설계하셨습니다.  이번 과제의 데이터는 엑셀 형식이었고, 주제는 영업 및 마케팅 부서직원으로써 가상회사의 판매 데이터(고객, 상품, 상품판매, 주문정보, 판매채널, 날짜 등의 정보로 구성)를 분석-가공-시각화-예측-처방 등의 단계를 거쳐 잠재 고객 발굴, 이탈 고객 방지, 우수 고객 재구매 촉진, 혜택 제공 등의 효과적인 마케팅 방안을 도출하는 과제였습니다.  처음 데이터를 접했을 때, 7개의 시트와 각각 수천에서 수십만 개의 데이터가 존재한다는 점에서 복잡성과 방대한 작업량에 압도감을 느꼈습니다. 특히, 시트 간의 열 구조가 다르고 데이터가 연결되지 않은 상태였기 때문에 데이터를 병합하고 분석 가능한 형태로 만드는 데 상당한 어려움이 있을 것이라 생각되었습니다. 이러한 문제는 작업의 출발점에서부터 많은 시간과 노력이 요구되는 도전 과제였습니다.  효과적인 마케팅 방안을 도출하기 위해서는 방대한 데이터를 빠르고 정확하게 처리할 필요가 있었고,  데이터 분석의 과정에서 복잡한 코드를 작성하거나 시각화 아이디어를 도출하는 데 많은 시간이 소요될 것으로 예상되었지만 이 과정에서 ChatGPT를 활용하면 데이터 분석 프로세스를 효율적으로 최적화할 수 있을 것이라고 판단하였습니다.  분석 아이디어를 도출하고 Python 기반의 분석 코드와 시각화 코드를 생성하는 등 ChatGPT를 적극적으로 활용하여 작업 시간을 획기적으로 단축할 수 있었습니다. 데이터 병합 작업은 절약된 시간을 통해 데이터를 정리하고 시각화하여 과제를 보다 신속히 완성할 수 있었습니다.  최종적으로 분석 결과를 PDF로 정리해 제출하는 과정에서도 ChatGPT의 도움을 받아 문서의 논리적 구성과 가독성을 높였습니다. 예를 들어, 분석 과정에서 도출된 여러 그래프와 수치 데이터를 이해하기 쉽게 정리할 수 있도록 ChatGPT가 요약문을 생성하고, 각 항목 간의 연결성을 높이기 위한 적절한 제목과 소제목을 추천해 주었습니다. 또한, 결론부 작성에서는 주요 인사이트를 간결하게 정리하며 설득력 있는 문장을 완성하는 데 큰 도움을 받았습니다. | |
| **5. 사용 프롬프트** | { 공백 포함 300자 이상 작성 }  프롬프트 구성, 활용방식, 입력값, 출력값 나누어 자세하게 작성해주세요. |
| **1) 분석 아이디어 도출**  **프롬프트: "대학교 데이터 분석 과제에서 사용할 흥미로운 분석 주제를 10가지 제안해주세요."**  **입력값: 데이터 시트명과 컬럼 정보를 제공**  **출력값: 고객 세분화, 매출 기여도 분석, 인기 상품 분석 등 다양한 분석 주제 제안**  **2) 데이터 처리 및 분석 코드 생성**  **프롬프트: "다음 데이터 구조를 사용해 Recency, Frequency, Monetary 분석을 위한 Python 코드를 작성해줘."**  **입력값: 데이터의 예시 구조 및 열 이름**  **출력값: RFM 분석에 필요한 데이터 병합, 계산, 시각화 코드**  **프롬프트: "AdventureWorks 데이터에서 시트별로 칼럼 이름을 나열하고 데이터 병합 코드를 작성해줘."**  **입력값: AdventureWorks 데이터의 전체 구조**  **출력값: 칼럼 이름 목록과 병합에 필요한 코드 생성**  **프롬프트: "AdventureWorks 데이터를 기반으로 상위 5개 제품의 매출 기여도를 파악하는 파이썬 코드를 작성해줘."**  **입력값: 제품 및 판매 데이터 예시**  **출력값: 매출 기여도를 계산하는 데이터 처리 및 분석 코드**  **3) 시각화 코드 생성**  **프롬프트: "상위 5개 제품의 매출 기여도를 파이 차트로 시각화하는 코드를 작성해줘."**  **입력값: 상위 매출 제품 데이터 예시**  **출력값: Matplotlib를 활용한 파이 차트 시각화 코드**  **프롬프트: "RFM 분석 결과를 히스토그램으로 시각화하는 코드를 작성해줘."**  **입력값: RFM 분석 결과 데이터**  **출력값: Recency, Frequency, Monetary 별 히스토그램 시각화 코드**  **프롬프트: "AdventureWorks 데이터에서 지역별 매출 데이터를 시각화할 바 차트 코드를 작성해줘."**  **입력값: 지역 및 매출 데이터 예시**  **출력값: Matplotlib를 활용한 바 차트 시각화 코드**  **4) 최종 문서 작성 보조**  **프롬프트: "데이터 분석 결과를 기반으로 최종 보고서 결론문을 작성해줘."**  **입력값: 분석 결과 데이터 및 주요 인사이트**  **출력값: 보고서 결론문 초안**  **프롬프트: "AdventureWorks 데이터 분석 과정과 주요 결과를 요약하는 보고서 서론을 작성해줘."**  **입력값: 데이터 분석 과정 및 결과**  **출력값: 보고서 서론 초안**  **프롬프트: "데이터 분석 과정에서의 시행착오와 개선 방안을 구체적으로 작성해줘."**  **입력값: 시행착오 사례와 해결 방법**  **출력값: 시행착오와 개선 방안을 정리한 내용** | |
| **6. 시행착오와 해결방안** | { 공백 포함 300자 이상 작성 }  원하는 결과를 도출하기까지 겪은 시행착오, 보완/개선 방법을 자세하게 작성해주세요. |
| 1. 초기 데이터 병합 과정의 오류  시행착오:  AdventureWorks 데이터에는 7개의 시트가 포함되어 있으며, 각 시트의 열 구조가 상이했습니다. 예를 들어, Sales\_data와 Product\_data를 병합할 때 ProductKey를 기준으로 병합해야 했지만, 데이터 구조를 정확히 이해하지 못해 병합이 실패했습니다. 데이터가 연결되지 않아 추가 분석 작업이 불가능했으며, 시트 간 데이터 매핑에서 누락된 정보가 다수 발생했습니다.  보완/개선 방법:  ChatGPT에 오류 메시지와 데이터를 입력하자, 병합의 실패 원인을 상세히 분석해 주었고, left\_on과 right\_on 옵션을 사용하는 병합 코드를 제공해 주었습니다. 이 코드는 데이터의 열 이름과 데이터 유형 간 불일치를 해결하는 방안까지 포함하고 있었습니다. 이를 통해 121,253개의 판매 데이터와 397개의 제품 데이터를 성공적으로 연결할 수 있었으며, 데이터 병합 과정에서 학습의 기회로 활용할 수 있었습니다.  2. 시각화 코드 실행 오류  시행착오:  상위 5개 제품의 매출 기여도를 파이 차트로 시각화하려는 시도에서, 데이터 정렬 및 전처리가 충분하지 않아 열(Sales Amount, Product)이 누락되었고, 결과적으로 시각화가 실패했습니다. 데이터가 정렬되지 않아, 차트 생성 시 올바른 레이블이 표시되지 않는 문제도 발생했습니다.  보완/개선 방법:  ChatGPT는 문제 원인을 정확히 지적하며, 누락된 열을 생성하기 위한 데이터를 준비하는 코드와 적절한 시각화 코드를 생성해 주었습니다. 이를 통해 매출 상위 5개 제품 데이터를 시각화하여 Mountain-200 시리즈의 Black, Silver 색상이 주된 매출원임을 도출할 수 있었습니다. 시각화 과정에서 데이터 정렬과 필터링의 중요성을 깨닫게 되었고, 이를 반복적으로 검토하는 습관을 들이게 되었습니다.  3. 결론 도출의 방향성 부족  시행착오:  데이터 분석 결과를 요약하고 결론을 작성하는 과정에서, 방대한 데이터로 인해 초점이 흐려졌습니다. 무엇을 강조해야 할지 명확하지 않아, 보고서 작성 시간이 크게 지연되었습니다. 주요 인사이트를 효과적으로 전달하지 못할 우려가 있었습니다.  보완/개선 방법:  ChatGPT에 분석 데이터를 제공하며 "주요 인사이트를 3~5가지로 요약해줘"라는 프롬프트를 입력하니, 데이터의 핵심을 빠르게 도출할 수 있었습니다. 특히, 고객 세그먼트화, 매출 상위 제품 기여도, 그리고 지역별 판매 성과와 같은 주요 인사이트를 정리해 주었습니다. 이를 기반으로 논리적이고 설득력 있는 결론을 완성할 수 있었으며, 보고서 작성 시간도 크게 단축되었습니다.  4. 데이터 크기 및 복잡성 처리  시행착오:  AdventureWorks 데이터는 총 18485명의 고객 데이터, 121,253개의 판매 데이터, 397개의 제품 데이터 등 방대한 양으로 구성되어 있었습니다. 데이터가 지나치게 방대하여 수작업으로 처리하기 어려웠고, 데이터의 구조를 파악하는 데만 많은 시간이 소요되었습니다.  보완/개선 방법:  ChatGPT의 도움으로 필요한 데이터를 정확히 필터링하고, 고객 세그먼트화 및 매출 기여도 분석 같은 구체적인 작업을 신속히 수행할 수 있었습니다. 반복적이고 시간이 많이 소요되는 작업을 자동화하여, 분석 과정에서의 부담을 크게 줄일 수 있었습니다. 이 과정에서, ChatGPT가 제공한 코드는 데이터 전처리와 분석의 흐름을 이해하는 데 중요한 학습 자료로 활용되었습니다. | |
| **7. 생산성 향상 정도** | { 공백 포함 300자 이상 작성 }  Before/After 나누어 생산성 개선정도를 구체적으로 작성해주세요. |
| Before:  데이터 크기 및 복잡성: 총 7개의 시트와 각각 수천~수십만 개의 데이터를 수작업으로 병합하고 정리하는 데 평균적으로 8~10시간 소요. 시트 간 열 구조가 다르고 데이터가 방대해, 병합 오류가 자주 발생.  분석 주제 도출: 적절한 분석 방향을 설정하는 데 평균 2~3시간 필요. 창의적이고 흥미로운 주제를 찾는 데 어려움.  Python 코드 작성: 데이터 전처리 및 시각화 코드를 직접 작성하며 디버깅하는 데 3~5시간 소요. 데이터 구조에 익숙하지 않아 오류가 반복적으로 발생.  최종 보고서 작성: 결과물을 정리하고 논리적으로 구성하는 데 4~5시간 소요. 결론 도출 및 서술에 시간이 많이 소요.  After:  데이터 병합 및 정리: ChatGPT의 도움으로 데이터 병합 코드를 자동 생성하여 2시간 이내로 단축. ChatGPT는 데이터 병합의 논리를 설명하며, 단순한 작업 이상의 학습 효과를 제공.  분석 주제 도출: ChatGPT로 10분 이내에 흥미로운 분석 주제를 여러 개 생성. 특히 고객 세분화, 매출 기여도 분석, 인기 상품 분석 등 다각적이고 실질적인 아이디어 제공.  코드 작성 및 디버깅: 복잡한 Python 코드를 작성하고 오류를 수정하는 시간을 50% 이상 단축. ChatGPT는 코드를 단계적으로 설명하며, 사용자에게 작성한 코드의 맥락과 의미를 이해시키는 데 도움.  보고서 작성: ChatGPT의 문서 보조 기능을 활용해 최종 보고서를 1시간 이내에 완성. 주요 인사이트를 간결하게 요약하며, 설득력 있는 결론 작성 가능.  구체적인 생산성 향상 효과:  전체 작업 시간은 기존 20시간 이상에서 10시간 이내로 대폭 단축.  121,253개의 판매 데이터를 빠르게 병합하고, 18,485명의 고객 데이터를 분석하여 RFM 세그먼트를 생성하는 데 걸리는 시간이 8시간에서 2시간으로 감소.  Python 코드를 작성하고 디버깅하는 데 필요한 시간은 기존 3~5시간에서 1시간으로 단축되었으며, 오류 발생률 또한 크게 감소.  데이터 시각화 작업은 기존 2시간 이상 소요되던 작업이 30분 이내로 줄어듦.  최종 보고서 작성 시간은 기존 4~5시간에서 1시간 이내로 감소하며, 보고서의 완성도와 논리성이 크게 향상.  ChatGPT를 통해, 복잡하고 부담스럽던 데이터 분석 과제가 오히려 창의적이고 즐거운 도전 과제가 되었습니다. 반복적인 작업이 자동화됨으로써 더 많은 시간을 인사이트 도출과 시각화에 투자할 수 있었으며, 분석 과정에서의 실수를 줄이며 학습 경험을 극대화할 수 있었습니다. ChatGPT는 단순한 도구가 아니라, 나만의 10년차 데이터 분석가로 느껴졌습니다. | |

## III. 공모작 세부 [선택 작성]

|  |  |
| --- | --- |
| **[선택1]**  **사용한 실제 프롬프트 링크** | { 분량 자유 }  필수 4번 항목과 연계되는 추가 항목으로, 제출 시 심사에 더 도움이 됩니다. |
| https://chatgpt.com/share/675423b6-d084-8003-8f8c-22daa859d0d2 | |
| **[선택 2]**  **나만의 활용 노하우** | { 분량 자유 }  공모작과 관련한 재미있는 에피소드가 있다면 작성해주세요. |
| **ChatGPT에게 "질문 폭격" 하기**  **제 나름의 노하우는 ChatGPT에게 질문을 끊임없이 던지며 점진적으로 문제를 해결하는 "질문 폭격" 전략입니다. 한 번에 완벽한 답변을 기대하기보다, 여러 번 질문을 세분화해 답변의 정확성과 세부성을 높이는 방식입니다. 예를 들어, 처음에는 "상위 5개 제품의 매출 기여도를 분석해줘"라는 단순한 질문으로 시작했습니다. 이후 "매출 데이터가 누락된 경우 어떻게 처리해야 하나요?", "파이 차트의 레이블이 겹치지 않도록 조정할 방법은?"과 같은 세부 질문을 던졌고, 이를 통해 데이터 전처리와 시각화 과정을 완성도 높게 마칠 수 있었습니다.**  **또한, ChatGPT의 답변이 이해되지 않을 때는 "이 코드를 한 줄씩 설명해줘"라고 요청했습니다. 이 과정에서 단순히 결과물을 얻는 것뿐만 아니라, Python 코딩 실력도 함께 향상되는 효과를 얻을 수 있었습니다.** | |
| **[선택 3]**  **보조 AI 프로그램 정보** | { 분량 자유 }  ChatGPT 외에 보조적으로 사용한 AI 프로그램이 있다면 작성해주세요. |
|  | |

## IV. 별도 제출 [필수작성]

|  |  |
| --- | --- |
| **8. 구동 영상 URL (실행화면)** | 실제로 생활 혹은 업무에 활용한 과정을 녹화하여 아래 두 가지 방법 중 한 가지를 택하여 제출해주세요.  **· [택1]** 녹화된 파일을 확인할 수 있는 (다운로드 가능한) URL을 아래에 입력 　　　 ( 유튜브 링크로 제출 시 보안을 위해 일부 공개로 설정 필수 한 뒤 제출해주세요. )  **· [택2]** 아래 칸에는 파일로 제출이라고 작성한 뒤, 파일명을 다음과 같이 지정하여 메일로 첨부   ( Track1\_개인/팀\_패캐머\_별도제출영상 ) |
| 파일로제출(용량제한으로 촬영 분량의 절반(20분가량) 밖에 못 찍었습니다..) | |