

웹 개발 1차 / 프롬프트 엔지니어링

날짜	@February 14, 2024
태그	

▼ 전처리 코드

▼ def excel_preprocess 함수로 전처리 진행시 통합 데이터프레임 반환

```
def date_data(xlsxfile, sheet_no):
    a=pd.read_excel(xlsxfile, sheet_no)
    a.dropna(axis=0, how='all', inplace=True)
    a.dropna(axis=1, how='all', inplace=True)
    a.dropna(axis=0, how='any', inplace=True)

    hl=list(np.where([x=='행 레이블' for x in list(a.iloc[:,0])])[0])
    zz=a.iloc[hl[-1]:]
    al1=list(np.where([type(x)==str for x in list(zz.iloc[:,0])])[0])
    c=[]
    for i in al1:
        if i==0:
            cl=zz.iloc[i]
        else:
            cl2=zz.iloc[ind+1:i]
            if len(cl2)>=2:
                cl2.columns=cl
                cname=zz.iloc[ind,0]
                cl2['sort']=cname
                cl2["예금+대출율"] = cl2["예금+대출율"].replace('%', '')
                cl2["예금+대출율"] = cl2["예금+대출율"].astype(float)
                cl2["클릭율"] = cl2["클릭율"].replace('%', '')
                cl2["클릭율"] = cl2["클릭율"].astype(float)
                c.append(cl2)

            ind=i
    z=pd.concat(c)
    return(z)

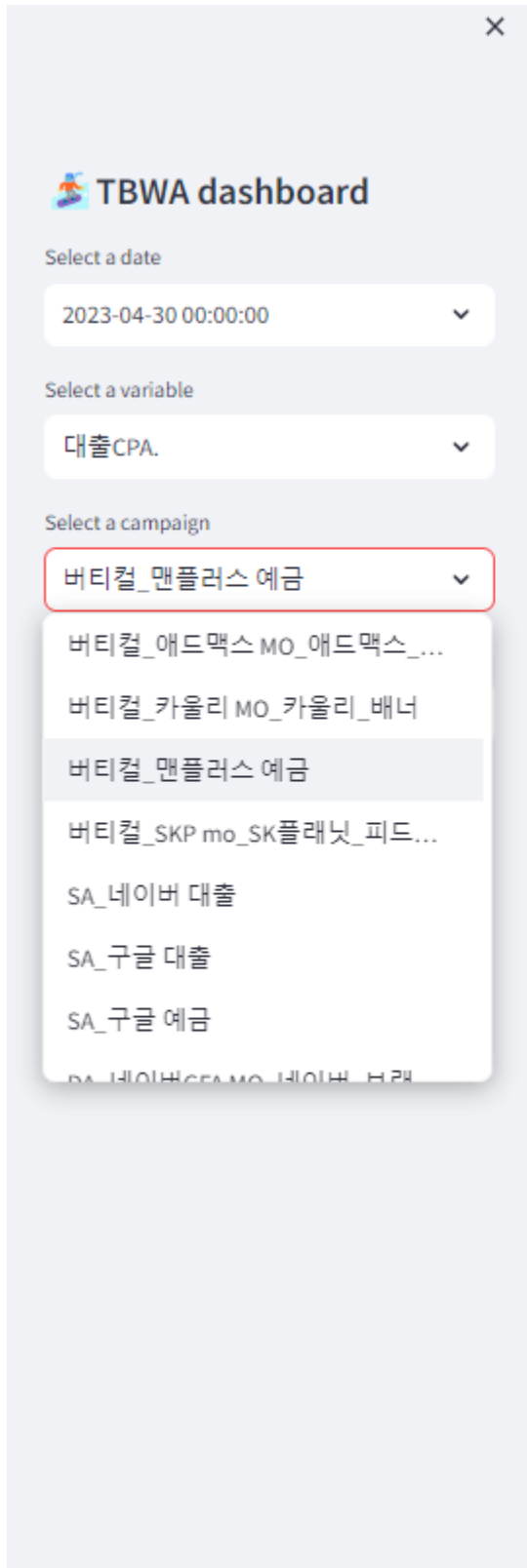
# 엑셀파일 입력시 단일 df 반환, 'talbe_sort' 컬럼에서 세부 캠페인 확인 가능
def excel_preprocess(exlfile):
    sheetnm=pd.read_excel(exlfile, None)

    sheetn=list(sheetnm.keys())
    sheetn.pop(4)
    sheetn.pop()
    sheetn.pop(0)

    stlist=[]

    for i in sheetn:
        ff=date_data(exlfile, i)
        ff['sort']=i+' '+ff['sort']
        stlist.append(ff)
    xlsx_sum=pd.concat(stlist)
    return(xlsx_sum)
```

▼ streamlit 라이브러리 구현 대시보드



▼ 사이드바, 선택상자 코드

```
with st.sidebar:
    st.title('🏠 TBWA dashboard')

    year_list = list(Tbdata['행 레이블'].unique()[::-1])
    var_list = list(Tbdata.columns)[::-1]
    com_list = list(Tbdata['sort'].unique()[::-1])
    selected_year = st.selectbox('Select a date', year_list)
    selected_var = st.selectbox('Select a variable', var_list)
    selected_com = st.selectbox('Select a campaign', com_list)

    df_selected_year = Tbdata[Tbdata['행 레이블'] == selected_year]
    df_selected_year_sorted = df_selected_year.sort_values(by=selected_var, ascending=False)

    color_theme_list = ['blues', 'cividis', 'greens', 'inferno', 'magma', 'plasma', 'rocket']
    selected_color_theme = st.selectbox('Select a color theme', color_theme_list)
```

▼ 대시보드 링크

- <https://strsample-2fhq3rthvvbzane5rnyuuz.streamlit.app/> 수정한 대시보드


- <https://github.com/dataprofessor/population-dashboard> 깃허브 참고하여 대시보드 수정했음
- <https://strsample-waaqpnxxaks47rw33fdj92.streamlit.app/> 2/13 기준
 - 구현 상황:
 - 데이터 입력 기능
 - 엑셀 데이터 전처리 기능
 - 특정 날짜 구간 지정 및 리스트 반환 기능
 - 특정 캠페인 및 세부종목 지정 기능
 - 캠페인 선 그래프 출력 기능(데이터 정리가 쉽게 가능하므로 다양한 시각화 진행에는 문제 없음)
 - 구현 필요 요소:
 - 각 지표별 연산식 필요
 - KPI 달성 관련 데이터 추출 기능
 - 전일 비교 기능
 - 코멘트 작성 기능 + 코멘트 작성에 필요한 정보 정리 기능

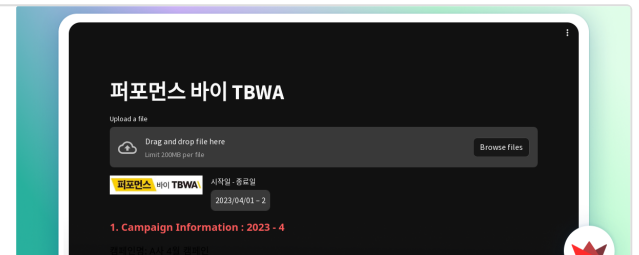


웹 개발 진행 상황

퍼포먼스 바이 TBWA

A dashboard web app template built in Python using Streamlit.

 <https://strsample-waaqpnxxaks47rw33fdj92.streamlit.app/>



[tbwa_app.py](#)



프롬프트 엔지니어링 진행 상황

모델 이슈

- 라마 설치 시도 아직 해결 안됨.
- 번역기 사용 → 번역기 성능, 비용, 글자 수 등 고려해야 할 사항이 많음.
- 라마 보단 ChatGPT 사용이 나을 것 같다고 판단이 되며 현재는 GPT 3.5와 루틴 GPT 4.0으로 시도 중

결론 : 번역기 사용에 대한 해결책이 나오면 라마 사용에 대해서 고려할 예정. 하지만 영어, 한국어 모두 GPT가 더 성능이 우수한걸로 판단됨.

메일 답변

먼저, 광고 상품에 대해 현 일을 기준으로 일주일 기간 동안의 데이터 값을 대시보드에 보여주고 전일 대비 코멘트 요약 생성을 목표로 하려고 합니다. 웹페이지를 기준으로 코멘트가 생성되어야 하기 때문에 매체별로 일자 비교에 대한 프롬프트 고도화에 집중하려고 합니다. 이후 월별 또는 매체별 비교에 대한 프롬프트를 작성하려고 합니다.

>> 네, 좋습니다. 가장 많이 활용하는게 데일리 / 먼슬리 리포트이기 때문에 일별과 월별 비교는 좋은 생각입니다.

매체별 비교에 대한 부분은 예를 들면 구글과 크리에오를 비교하는 것을 의미하는 걸까요?단순히 매체끼리의 비교는 유의미 하지 않을 것 같습니다.

매체 안에서도 상품마다 광고 영역이 다르고, 목적과 특성이 달라서

"지속적으로 카카오DA가 낮게 나오면 이전 추세에 비해 최근 광고 효율이 떨어지니" 소재를 개선하거나, 다른 전략을 사용해보자는 의견을 줄 수 있지만, '어제 구글DA에 비해 카카오 DA효율이 낮으니 캠페인 운영성과가 좋지 않다'고 단순히 판단하기는 어려움이 있을 것 같습니다.

조금 더 발전해본다면, 긍정적인 변화가 나타났을 때 특정 캠페인- 광고그룹 - 광고소재 중에서 기존보다 변화가 더 많은 항목이 있다면 그걸 노티해주는게 더 좋을 것 같습니다.

두번째로 코멘트를 작성할 때 증감율 3%를 기준으로 프롬프트를 생성하려고 합니다. 지난 미팅 때 튀는 값에 대해 말씀해주셨었는데 그에 대한 기준이 모호하여 전일 대비 3% 이상을 긍정적인 변화라고 가정하려고 합니다.

>> 네, 해당 기준은 광고주별로 달라질 수 있겠지만, 해당 부분은 케이스에 따라 변동이 가능한 기준으로

현재 임의의 광고주 A사에 대한 가이드라인을 3%로 설정하는 것에 대해서는 이견 없습니다.

세번째로 프롬프트의 결과에 대한 성실도, 만족도를 체크하기 위해서 체크리스트를 만들었습니다. 체크리스트의 사항이 적절한지, 또 추가 사항이 있는지 궁금합니다.

[체크리스트]

- 3% 이상의 변화를 보인 것에 대해서 언급하고 있다.
- [] 변화에 대해서 다른 지표와의 상관관계를 평가하고 있다.
- [] 긍정적인 지표만 언급하고 있다.
- [] 사실 나열 이외에 평가적인 요소가 들어간다.
- [] 작성된 답변에 있는 숫자가 정확하다. (할루시네이션이 없다.)

>> 대체로 들어가야 할 항목이 들어가있는 것 같습니다.

- 코멘트에서 잘 쓰지 않는 단어나 어색한 문장에 대한 교정

마지막으로 결과 형식에 대한 것입니다. [분석 결과]와 [캠페인에 대한 평가] 항목으로 나누어서 결과를 받으려고 합니다. 아래와 같은 형식이 적절한지 궁금합니다.

[분석 결과]

- CPC는 14.69% 감소하여 91으로 나타났습니다. 이는 긍정적인 변화로 판단됩니다.
- CPU는 7.13% 감소하여 1,794,029로 나타났습니다. 캠페인의 효율이 향상되었음을 나타냅니다.
- 신규방문CPU는 3.21% 감소하여 316으로 나타났습니다. 이는 어느 정도 감소했지만, 여전히 안정적인 수치입니다.
- 심사CPA는 30.35% 감소하여 448,507로 나타났습니다. 이는 매우 큰 감소로 효율적인 캠페인 운영을 시사합니다.

[캠페인에 대한 평가]

1. 캠페인의 효율이 향상되었음을 확인할 수 있습니다. CPC와 CPU가 감소하면서 더 많은 클릭을 더 적은 비용으로 확보하고 있습니다.
2. 특히, CPC의 14.69%의 큰 감소는 광고 클릭을 얻는 데 더 효율적으로 예산을 사용하고 있다는 것을 시사합니다.
3. CPU의 감소는 방문자수와 클릭 수의 감소와 관련이 있을 것으로 판단됩니다. 하지만, 이는 신규방문CPU의 감소로 크게 영향받지 않고 있다는 것을 의미합니다.
4. 신규방문CPU와 심사CPA의 감소는 캠페인이 더 효과적으로 신규 방문자를 유입시키고, 심사 단계에서도 더 효율적으로 비용을 소모하고 있다는 것을 나타냅니다.

종합적으로, 현재 캠페인은 효율적으로 운영 중이며, 특히 CPC와 CPU의 감소는 예산을 효율적으로 활용하고 있음을 시사합니다. 신규방문과 심사 관련 지표도 안정적으로 유지되고 있으므로 현재 전반적으로 긍정적인 결과를 보이고 있습니다.

이상으로 지금까지 진행한 프롬프트 결과입니다.

>> 대시보드 샘플 이미지로는 확인이 어려워서요.

분석 결과나 평가에서 특정 매체 / 특정 매체의 특정 캠페인CPC가 아닌 전체의 CPC를 보고 있는 것 같아 문의드립니다.

1) 일별 / 월별 기준 설정 >> 데일리 리포트와 먼슬리 리포트를 선택하게 하여 이에 따라 다른 프롬프트를 제공해도 되지만, 어렵다면 데일리 리포트 사이트와 먼슬리리포트 사이트를 분리해도 됩니다.

2) 매체별로 성과 코멘트 작성

3) 캠페인 전체에 대한 성과 코멘트 작성

결과에 대한 코멘트 자체의 형식은 괜찮은 것 같습니다.

마지막 질문에서 말씀해주신 것 중 한 가지 궁금한 것이 있습니다.

캠페인 전체에 대한 성과 코멘트 작성은 보내주신 데이터 자료 중 summary total을 참고하여 작성하면 되는 것인지 궁금합니다.

매체별 코멘트는 각 매체별 시트의 daily trend를 참고하여 작성하였습니다.

>>

네, 전체 내용은 summary 기준으로 (summary – 각 매체 토탈값)

매체 별은 매체 별 이름 시트에서 확인해주시면 됩니다.

2. 프롬프트 시도

- 몇 가지 단어들 변경 (예, 제약 조건 → 꼭 지켜야할 규칙 등) → 성능이 조금 좋아짐
- 출력문에서 구체적인 예시 삭제
- ‘긍정적인 것’ → ‘-3% 이상 인 것’ 등으로 변경

1. 프롬프트

명령문

당신은 광고 퍼포먼스 마케터입니다. 광고 캠페인에 대하여 데일리 리포트 코멘트를 작성할 것입니다. 아래의 특징을 가지고 리포트 코멘트를 작성합니다. 증감을 데이터로 포함하여 비용 관련 지표를 중점적으로 살펴봅니다.비용과 관련된 지표는 감소하면 긍정적인 변화입니다.

코멘트 작성 시 꼭 지켜야 할 규칙

코드블록을 사용하지 않는다.긍정적인 변화에 대해서만 코멘트를 작성한다.증감율에서 3% 이상의 변화가 있는 “모든” 지표에 대해서만 코멘트를 작성한다.

비용 관련 지표

CPC, CPS, CPU, 신규방문CPU, 접수CPA, 심사CPA, 승인 CPA, CPA, 예금CPA, 대출CPA

코멘트 출력문 예시

[분석 결과]

CPC는 xx% 감소하여 xx으로 나타났습니다. 이는 긍정적인 변화로 판단됩니다.CPU는 xx% 감소하여 xx로 나타났습니다. 캠페인의 효율이 향상되었음을 확인할 수 있습니다.

[캠페인에 대한 평가]

캠페인의 효율이 향상되었음을 확인할 수 있습니다. CPC와 CPU가 감소하면서 더 많은 클릭을 더 적은 비용으로 확보하고 있습니다. 이는 캠페인이 더 효과적임을 시사합니다.

2. 결과

[분석 결과]

- CPC는 14.69% 감소하여 91로 나타났습니다. 이는 광고 클릭당 비용이 줄어든 것으로, 긍정적인 변화로 판단됩니다.
- 광고비(VAT별도)는 7.13% 감소하여 1,794,029로 나타났습니다. 이는 총 광고비용의 줄어듦을 의미하며, 긍정적인 변화로 판단됩니다.
- 신규방문CPU는 3.21% 감소하여 316으로 나타났습니다. 이는 신규 방문자당 비용이 감소한 것으로, 캠페인의 효율이 향상되었음을 시사합니다.
- **심사CPA는 18.52% 감소하여 3,136으로 나타났습니다. 이는 심사 수당 비용의 큰 감소로, 효율적인 캠페인 운영을 나타냅니다.
- 승인CPA는 23.80% 감소하여 46,001로 나타났습니다. 이는 승인을 얻는데 드는 비용이 줄어든 것으로, 효율적인 캠페인 운영을 나타냅니다.
- CPA는 38.09% 감소하여 299,005로 나타났습니다. 이는 전체 캠페인 비용 효율성이 크게 향상된 것을 의미합니다.
- 예금CPA는 53.57% 감소하여 897,015로 나타났습니다. 이는 예금을 얻는데 드는 비용이 크게 줄어든 것으로, 예금 유치에 대한 비용 효율성이 향상된 것을 의미합니다.
- 대출CPA는 30.35% 감소하여 448,507로 나타났습니다. 이는 대출에 대한 비용 효율성이 향상된 것을 의미합니다.

[캠페인에 대한 평가]

1. 캠페인의 효율이 전반적으로 향상되었음을 확인할 수 있습니다. CPC, 신규방문CPU, 심사CPA, 승인CPA, CPA, 예금CPA, 대출CPA가 모두 감소하여 캠페인의 효율성이 크게 향상되었음을 시사합니다.
2. 특히, CPA의 38.09%의 큰 감소는 캠페인 전체의 비용 효율성이 크게 향상되었음을 시사합니다. 또한, 예금CPA의 53.57%의 감소는 예금 유치에 대한 비용 효율성이 크게 향상되었음을 보여줍니다.
3. 신규방문CPU의 감소는 신규 방문자를 더 효과적으로 유입하면서도 비용을 절감하고 있다는 것을 의미합니다.
4. 심사CPA와 승인CPA의 감소는 심사 및 승인 단계에서도 비용 효율성이 향상되었음을 보여줍니다. 이는 캠페인이 더 효과적임을 시사합니다.

3. 체크리스트 6/7

- [v] 3% 이상의 변화를 보인 것에 대해서 언급하고 있다.
- [v] 변화에 대해서 다른 지표와의 상관관계를 평가하고 있다.
- [v] 긍정적인 지표만 언급하고 있다.
- [v] 사실 나열 이외에 평가적인 요소가 들어간다.
- [] 작성된 답변에 있는 숫자가 정확하다. (할루시네이션이 없다.) **→ 증가한 것에 대해서 감소했다고 평가함**
- [v] 코멘트에서 잘 쓰지 않는 단어나 어색한 문장에 대한 교정이 이루어지고 있다.
- [v] 출력문 형식을 지키고 있다.

4. 더 개선되어야 할 점

- 작성된 답변에 있는 숫자의 정확도 높이기
- 다른 지표와의 상관관계 평가에 대한 부분을 구체화시키는 방법 모색



앞으로 진행할 것

- 웹 개발 1차 마무리
- 프롬프트 고도화 (영어 → 한국어 시도, GPT4 API로 시도)