1. 이번 대회 관련 관심 있는 데이터 (정형 데이터, 비 정형 데이터 등)
2. 주로 사용하는 데이터 분석 방법 (데이터 전 처리, AI 모델 등)
3. 팀별 특색, 강점과 약점

* 팀별로 분석한, 공부한 내용..?

<https://dacon.io/competitions/official/235647/overview/description>

**기존 데이콘 심리 성향 예측 AI 경진대회**

: 심리학 테스트의 범주가 넓어짐에 따라 해당 영역의 데이터 분석 방법 탐구

→ 국가 선거 투표자/미투표자의 심리학적 성향을 분석

<https://www.dacon.io/competitions/official/235647/talkboard/401534?page=1&dtype=recent&ptype=pub>

→ 데이터 변수 설명

**기존 구현/연구/아이디어 살펴보기**

* 1위 → LGBM + NN Ensemble

1. 마키아벨리니즘 테스트 스코어 feature로 추가. (→ EDA 중요)
2. EDA 결과, 성별에 따라 Emotional Stability/ Conscience/ Open Minded가 투표 여부에 미치는 영향이 크다고 판단되어 feature를 추가함. → 이런 eda 결과를 찾아내는게 중요할듯.

정보 출처: <https://www.sciencedirect.com/science/article/abs/pii/S0261379413001613>

* 최종 정확도

model1 : RandomForest/ Lgbm/ GradientBoosting Ensemble

Model2: Lgbm Ensemble with different features

model3 : nn

→ Final Ensemble

* 이상치 처리

**데이터가 극단적으로 치우쳐져있을때 log를 취하자/ XGB 모델**

<https://www.dacon.io/competitions/official/235647/codeshare/1775?page=1&dtype=vote>

참고해보면 좋을듯.

* log 사용해봤지만 성능 개선이 미미하여 사용하지 않았다는 글도 있으므로 참고해서 한번 해보는식으로 하면 될거같다.

<https://www.kaggle.com/code/azminetoushikwasi/divorce-reasons-analysis-with-pca-shap-tsne-rfe/data>

* 이거도 예측 모델 이런건데 데이터셋 보는거 도움 될거같아서 주소 넣었습니다