TAXI FARES PREDICTION

Ahmed Haitham 454778
Irena Zimovska 414522
Jakub Bandurski 417911
Pearly Tantra 455722

BACKGROUND

Why Choose Taxi Fares Prediction

for Reproducible Research?

- Relevance and Practicality

- Complexity of Factors

- Opportunity for Improvement

- Transition to Python

- Structured Programming Approach

- GIT for Team Collaboration

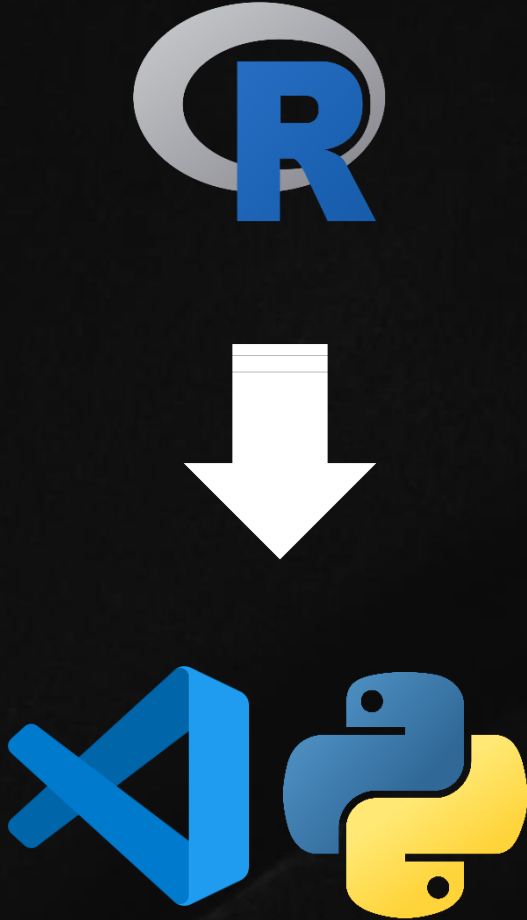- Focus on Model Optimization

- Overall Impact

BACKGROUND

how we want to solve the problem

- Enhanced ETL Process

- Structured Codebase

- GIT for collaboration

- Model Selection and Optimization

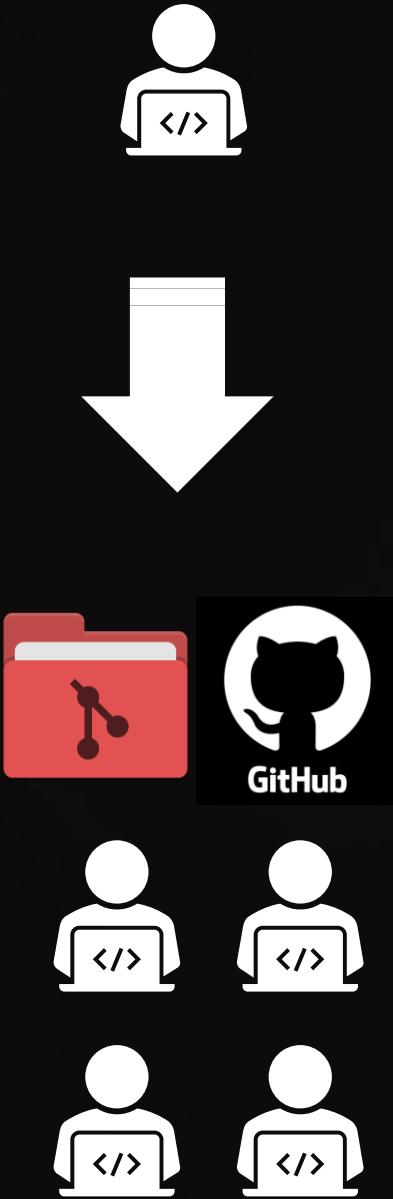- Validation and Testing
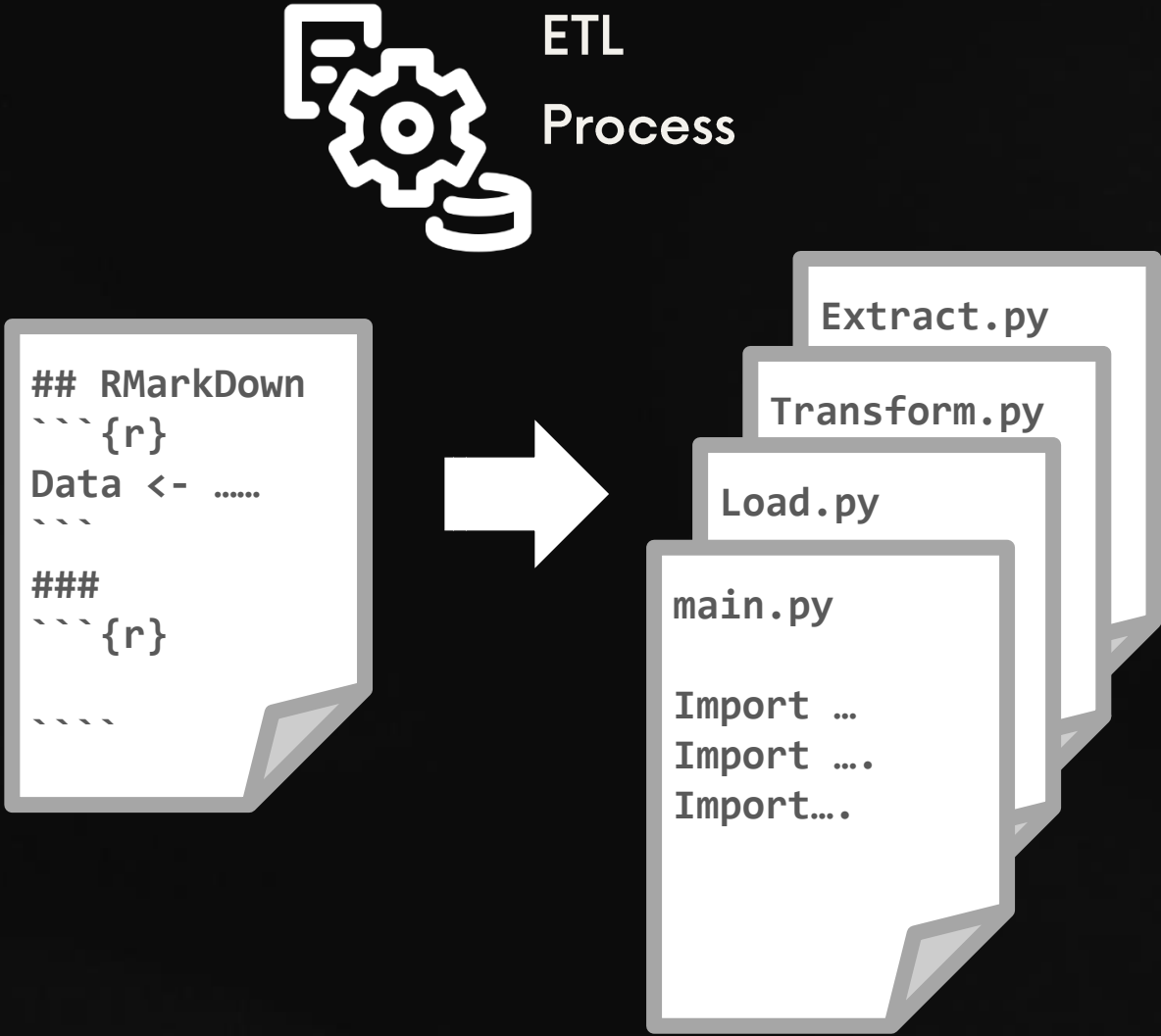
- Documentation and Reproducibility

TOOLS

Programming Language

Collaboration

Enhancement

ETL Process

```
## RMarkDown
```{r}
Data <- ……
```
###
```{r}

```
```

Extract.py

Transform.py

Load.py

```
main.py

Import …
Import ….
Import….
```

GitHub

TASK ALLOCATION

AHMED — Feature engineering + feature selection (more than 1 method)

IRENA — EDA + visualization

JAKUB — Modelling + Testing Pipeline

PEARLY — Clustering + Documentation

THANK YOU