OSS Project #2

12211618 박준용

Requirements

- `matplotlib`
- `numpy`
- `pandas`
- `python==3.9`
- `scikit-learn`
- `statsmodels`

Quick installation

```
pip install -r requirements.txt
```

Project #2-1 Data analysis with Pandas

1. Print the top 10 players in hits …
   - Implemented in function `print_players()`
2. Print the player with the highest war …
   - Implemented in function `print_highest_war_player_by_position()`.
3. Among R (득점), H (안타) …, which has the highest correlation with salary (연봉)?
   - Implemented in function `correlation()`.
   - Correlation was visualized using `matplotlib` and `plot_corr()` in `statsmodels` library.

Project #2-2 Data analysis with sklearn

- All requirements were implemented in a given function.
- In `calculate_RMSE()`, the `rmse()` function defined in the `statsmodels` library was used.
- A logic to store a model in the form of a binary has been added using `pickle`.