

WorkSense

Tufts Code for Good





Worksense x Code for Good



Worksense is a start-up launched by Tufts CS alumni Timi Dayo-Kayode and Kevin Destin. Their mission is to help companies understand how inclusive they are, and where to improve, across different metrics such as compensation gaps, hiring gaps, and more.

Our task was to use data science modeling to analyze people-data in the workplace and build a tool that can identify potential gaps in compensation between different demographics groups.

Data Collection

1	Gender	Race	Sexual orientation	Age	Average salary (weekly)	Position
2	Male	White	Heterosexual	16-19	1,721.21	SoftwareEngineer
3	Male	White	Heterosexual	16-19	1,380.24	Consultant
4	Male	White	Heterosexual	16-19	1,524.88	ProductManager
5	Male	White	Heterosexual	16-19	1,840.42	SoftwareEngineerManager

Methods

- Scraping online data sources including LinkedIn, levels.fyi, Glassdoor.
- Developed methodology to create synthetic dataset



Model Selection

Considerations:

- Interpretability of model results
- Flexibility of model input
 - Companies collect different personnel data
- Accuracy of model with minimal supervision

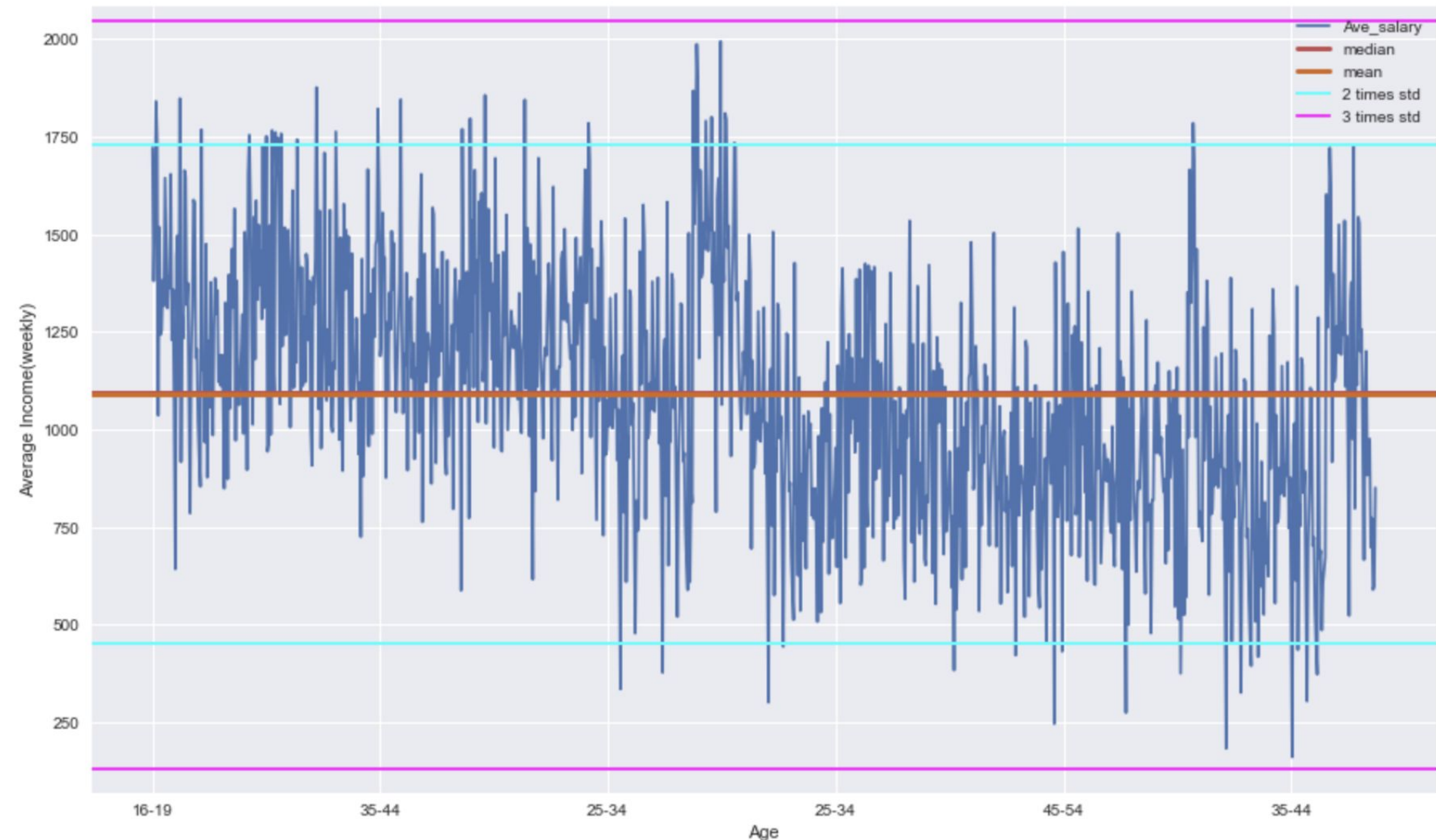
Models implemented:

- Linear Regression
- Logistic Regression
- Polynomial Regression
- Decision Trees
- Random Forest

We ultimately selected the Polynomial Regression as the primary model, and used Random Forests to cross-validate feature importance.

Final Model

Implemented a polynomial regression model and detected outlier compensation by standard deviation, while identifying relationships with variables that may cause bias



Average income distribution

```
-----  
Gender                Male  
Race                  Asian  
Sexual orientation    Heterosexual  
Age                   45-54  
Ave_salary            1994.04  
Name: 499, dtype: object  
903.8927627906976  
TOO HIGH  
-----
```



Impact

Worksense receives

- A working model that can be used to drive meaningful people analytics for their clients
- Well documented code so that they can further develop this code and implement it with their front end

Our lessons learned...