Natsman / **household-expenditure-analysis**

Code Issues Pull requests Actions Projects Wiki Security Insights Set

Household Expenditure Analysis using regression & decision tree models — includes notebook, PDF report, and visuals

0 stars 0 forks 0 watching Branches Activity
Tags

Public repository

1 Branch 0 Tags Go to file t Go to file Add file Code

Natsman Add files via upload 98de07c · now

2437038_Nataraj N (1).ipynb	Add files via upload	3 minutes ago
2437038_Nataraj N (3).html	Add files via upload	3 minutes ago
2437038_Nataraj N.pdf	Add files via upload	3 minutes ago
Inc_Exp_Data.csv	Add files via upload	now
README.md	README.md	1 minute ago
output_1_1.png	Add files via upload	3 minutes ago
output_1_11.png	Add files via upload	3 minutes ago
output_1_3.png	Add files via upload	3 minutes ago
output_1_4.png	Add files via upload	3 minutes ago
output_1_5.png	Add files via upload	3 minutes ago
output_1_6.png	Add files via upload	3 minutes ago
output_1_7.png	Add files via upload	3 minutes ago
output_1_9.png	Add files via upload	3 minutes ago

README

Household Expenditure Analysis

Author: Nataraj Narayan
Repo: Natsman/household-expenditure-analysis

📊 Project Overview

This project analyzes how **education, income, and demographic characteristics** affect household monthly expenditure.
Using a dataset of 50 households, I compare **Multiple Linear Regression** and **Decision Tree Regression** models to understand which approach best explains expenditure patterns.

The project is part of my MSc coursework in **Applied Economics + AI/ML**, and demonstrates how traditional econometrics and modern machine learning can complement each other.

🔑 Key Results

- **Income** and **household size** are strong positive predictors of monthly expenditure.
- **EMI/Rent** significantly increases outflows.
- **Number of earners** shows a negative relationship with expenditure (possibly due to savings behavior in multi-earner households).
- **Education** had a weak effect in this small dataset.
- Cross-validation shows **Linear Regression generalizes better** on this dataset, while Decision Tree captures non-linear splits but risks overfitting.

📂 Repository Contents

- `2437038_Nataraj.ipynb` — Jupyter Notebook with the full analysis, code, and visuals.
- `2437038_Nataraj.pdf` — Exported PDF report (ready-to-read version).
- `2437038_Nataraj.html` — HTML version of the notebook (optional, for GitHub Pages).
- `requirements.txt` — Python package dependencies.
- `.gitignore` — Ignoring cache/temporary files.
- `images/` — Preview images from the analysis (optional).

⚙️ How to Run

Clone the repository:

```
git clone https://github.com/Natsman/household-expenditure-analysis.git
cd household-expenditure-analysis
```

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

- **HTML** 63.9%
- **Jupyter Notebook** 36.1%

Suggested workflows

Based on your tech stack

SLSA Generic generator Configure

Generate SLSA3 provenance for your existing release workflows

Jekyll using Docker image Configure

Package a Jekyll site using the jekyll/builder Docker image.

More workflows Dismiss suggestions