Homework — Stage 09 Feature Engineering

Assignment

In the lecture, we learned how to create new features from raw variables.
Now, you will adapt those ideas to your own project data.

Tasks

1. Create **2–3 engineered features** in your dataset.
    - Each feature must include:
        - Implementation code
        - Markdown explanation of reasoning
        - Optional: visualization or correlation check
2. Save your work as either:
    - `notebooks/feature_engineering_<team>.ipynb`
    - OR `src/features.py`
3. Commit and push to your repo before the next session.

Step-by-Step Instructions

1. Open the starter notebook (`stage09_feature-engineering_homework-starter.ipynb`).
2. Replace the sample synthetic data with your project dataset.
3. Implement at least 2 new features.
4. Document each with a short rationale.
5. (Optional) Plot or test correlation with target variable.
6. Save and commit.

Grading Rubric (100 points)

- **20 points**: At least two features implemented
- **20 points**: Each feature documented with clear rationale
- **20 points**: Features connect to EDA insights
- **20 points**: Code is correct and reproducible
- **20 points**: Optional plots or tests (bonus for effort)

Example Submission Expectation

- Feature 1: `spend_to_income_ratio`
    - Code snippet: `df['spend_income_ratio'] = df['monthly_spend'] / df['income']`

    - Rationale: “Helps capture proportionality of spending to earning.”

- Feature 2: `rolling_spend_mean`
    - Code snippet: `df['rolling_spend_mean'] = df['monthly_spend'].rolling(3).mean()`
    - Rationale: “Captures 3-month trend in spending behavior.”
