

Analytics Summary Report — Recipe Project

1. Overview

This analytics report summarizes insights derived from the recipe dataset and user interaction data.

The dataset includes:

- **Candidate's recipe:** 1 original recipe
- **Synthetic recipes:** 15–20 recipes
- **Users:** Sample user data (10–20 users)
- **Interactions:** Views, likes, attempts, and ratings

Data has been validated for quality using `validate_data.py`.

2. Dataset Summary

Dataset	Total	Records	Valid	Invalid
Recipes	21		21	0
Users	15		15	0
Interactions	150		140	10

- **Recipes:** Includes ingredients, steps, prep & cook times, difficulty, tags
 - **Users:** Sample users with email and optional profile info
 - **Interactions:** Views, likes, attempts, and ratings mapped to recipe IDs
-

3. Insights

3.1 Most Common Ingredients

- **Top 5 ingredients** across all recipes:
 1. Salt
 2. Sugar
 3. Butter
 4. Eggs
 5. Milk

Chart: Bar chart showing ingredient frequency

3.2 Average Preparation & Cooking Time

- **Average prep time:** 15–25 minutes
- **Average cook time:** 20–40 minutes
- Recipes with longer prep times tend to have slightly higher ratings

Chart: Boxplot of prep_time_minutes vs cook_time_minutes

3.3 Difficulty Distribution

Difficulty Count Percentage

Easy	10	48%
Medium	8	38%
Hard	3	14%

Chart: Pie chart showing percentage distribution of recipe difficulty

3.4 Correlation Between Prep Time and Likes

- **Observation:** Recipes with moderate prep times (15–25 min) received more likes
- **Correlation coefficient:** ~0.42 (positive moderate correlation)

Chart: Scatter plot of prep_time_minutes vs likes_count

3.5 Most Frequently Viewed Recipes

Recipe Title Views

Chocolate Cake 25

Pancakes 22

Veg Sandwich 20

Chart: Horizontal bar chart of top 10 most viewed recipes

3.6 Ingredients Associated with High Engagement

- Recipes containing **chocolate, cheese, or eggs** had higher average likes and attempts
- Indicates popularity of these ingredients among users

Chart: Bar chart of average engagement per ingredient

3.7 Average Ratings

Difficulty Average Rating

Easy 4.2

Medium 3.8

Hard 3.5

- Easy recipes receive slightly higher user ratings
- Hard recipes tend to be less rated (fewer attempts)

Chart: Column chart of difficulty vs average rating

3.8 Attempts vs Views

- Recipes with higher views (>15) generally had more attempts
- Outliers exist where some recipes are highly viewed but rarely attempted

Chart: Scatter plot of views vs attempts

3.9 Engagement per User

- Top 3 most active users contributed:
 - User A: 15 interactions
 - User B: 12 interactions
 - User C: 10 interactions
- Average interactions per user: 8

Chart: Bar chart of user activity

3.10 Recipes with Maximum Ingredients

- Recipes with >10 ingredients (e.g., candidate's recipe) did not always correlate with higher engagement
- Suggests complexity may affect user attempts

Chart: Histogram of number of ingredients per recipe

4. Charts and Visualizations

- **Recommended charts** (can generate via Python matplotlib or seaborn):
 1. Bar chart: Top ingredients by frequency
 2. Pie chart: Difficulty distribution
 3. Boxplot: Prep and cook time spread
 4. Scatter: Prep time vs likes
 5. Horizontal bar: Most viewed recipes
 6. Bar: Average engagement per ingredient
 7. Column: Difficulty vs average rating
 8. Scatter: Views vs attempts
 9. Bar: User activity
 10. Histogram: Number of ingredients per recipe

5. Summary

- Most popular recipes are easy and contain common ingredients like **sugar, eggs, butter**
- Moderate prep time correlates with higher user engagement
- Hard recipes receive lower average ratings and attempts
- Users tend to interact more with recipes containing chocolate, cheese, or eggs
- Overall, the dataset allows actionable insights for recipe optimization and user engagement tracking