

Data Storytelling

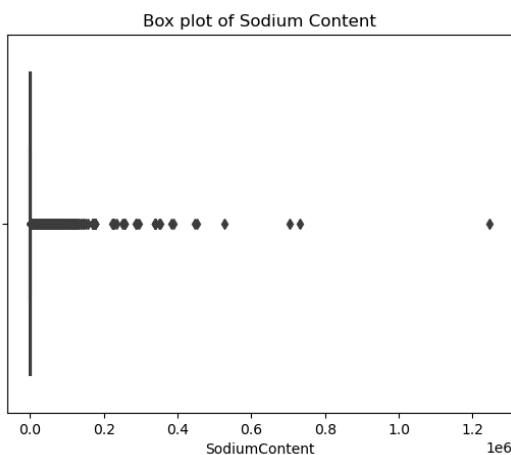
1. **Introduction:** This system is designed for everyone who opts for cooking at home, more beneficial for those with busy schedules. This system aims to utilize new recipes tailored to tastes and dietary preferences by accessing personalized suggestions to enhance the cooking experience.

2. **Data Collection:** The .csv file was accessed through kaggle. The key data points include recipe names, ingredients, cooking times, user ratings, and Nutritional information.

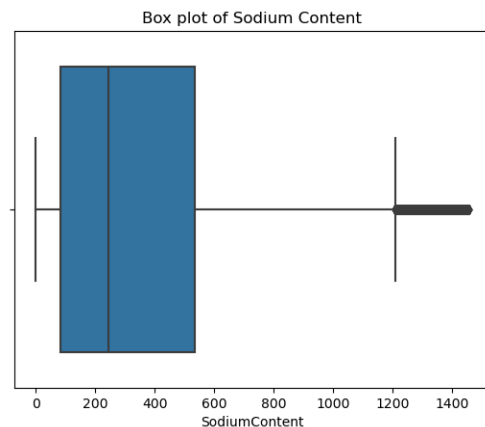
3. Data Analysis with visualization:

The first thing we did was identify outliers, and removing them, the following charts show the data before outliers were removed and after removal:

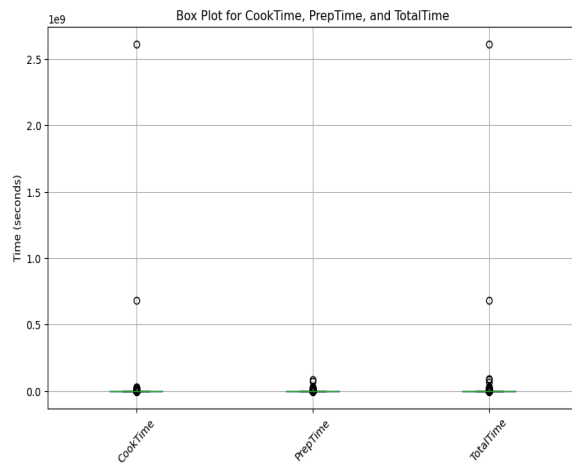
Original Sodium Content:



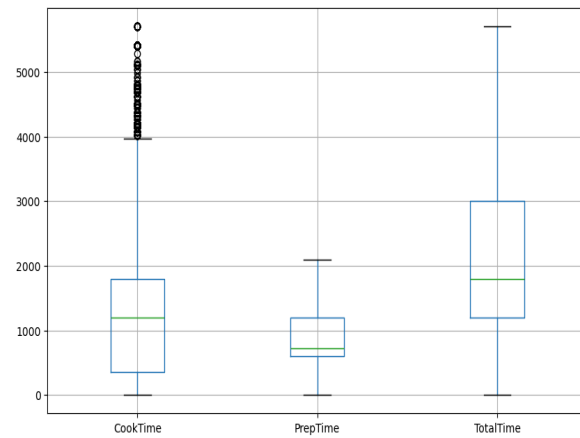
Sodium Content after Removing outliers::



Original Cooking Times:

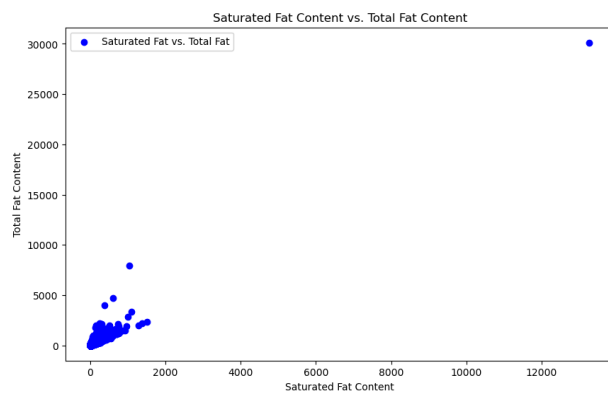


Cooking Times after Removing outliers:

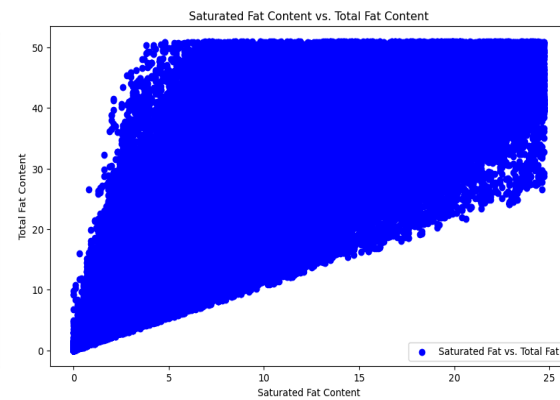


Relationship between TotalFat, Saturated Fat:

Original:



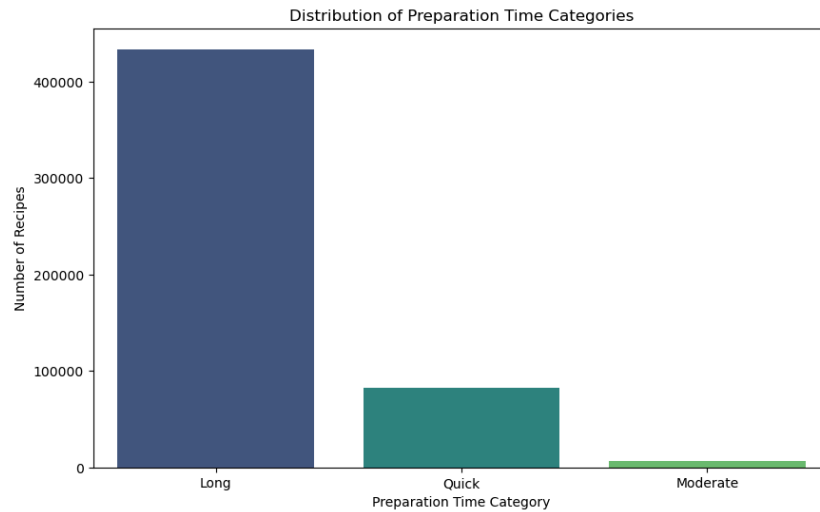
After removing outliers:



Techniques: Relationships between each category was analyzed.

No correlation between the nutritional properties vs. ratings was found.

The Distribution of CookingTime was visualized:



Aggregated Ratings vs. CookTime was evaluated, and no correlation was distinguished.

The following features were created:

1. An algorithm for nutritional features such as dairy-free, gluten-free, low-carb', 'vegetarian'.
2. Cuisine type was also categorized to: Italian, Mexican, Chinese, Indian, Japanese, Thai, Persian, American, Others.

```
•[69]: #Testing the feature:
       indian_recipes = df[df['CuisineType'] == 'Indian']
       print(indian_recipes)
```

	RecipeId	Name	AuthorId	CookTime	\
58	159	Chicken Curry	148316	2520.0	
94	236	Chicken Curry II	1597	1.0	
321	675	Curried Chicken II	1543	1.0	
605	2490	Cucumber Raita	1533	1.0	
684	2673	West Indian Bread Pudding	1549	1.0	
...	
324980	540873	Charishma's Gur Ki Roti With Khoya	6357	2100.0	
325020	540939	Chettinad Chicken	78626	1800.0	
325047	540977	INDIAN BASMATI With DRY FRUITS	2001004241	1800.0	
325143	541143	Instant Rava Dhokla Recipe	2002835253	120.0	
325260	541356	Masala Maggi Noodles in a Mug	2002835253	240.0	

	PrepTime	TotalTime	\
58	3900	6420	
94	3600	3600	
321	3300	3300	
605	0	0	

3. Categorizing easy vs. difficult recipes based on cooking time.

```
# Display the first few rows of the DataFrame to see the new column
print(df[['Name', 'TotalTime', 'Difficulty']].head())
```

	Name	TotalTime	Difficulty
0	Low-Fat Berry Blue Frozen Dessert	89100	Difficult
1	Cabbage Soup	3000	Medium
2	Buttermilk Pie With Gingersnap Crumb Crust	4800	Medium
3	A Jad - Cucumber Pickle	1500	Easy
4	Butter Pecan Cookies	3840	Medium

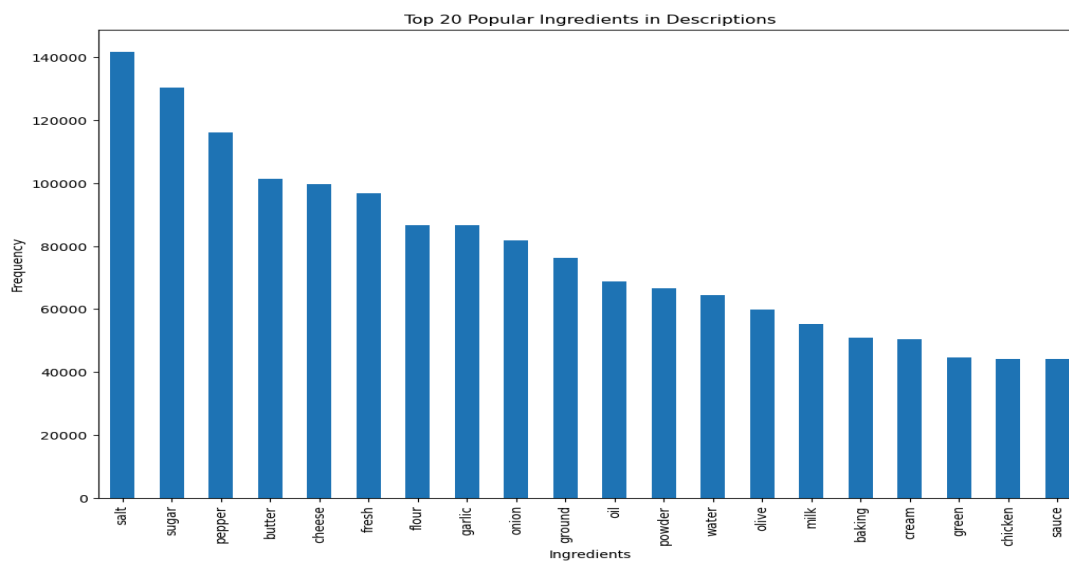
4. Categorized meat, seafood and vegetarian meals.

```
print(df[['Name', 'MealType']].head(-10))
```

	Name	MealType
0	Low-Fat Berry Blue Frozen Dessert	Vegetarian
1	Cabbage Soup	Vegetarian
2	Buttermilk Pie With Gingersnap Crumb Crust	Vegetarian
3	A Jad - Cucumber Pickle	Vegetarian
4	Butter Pecan Cookies	Vegetarian
...
325259	Chicken Pot Pie with Mashed Potato Crust	Meat
325260	Masala Maggi Noodles in a Mug	Vegetarian
325261	Chocolate Rum Snowballs	Vegetarian
325262	Cookie Cutter Shortbread Hearts	Vegetarian
325263	11-Minute Microwave Caramels	Vegetarian

[325264 rows x 2 columns]

5. 20 popular ingredients:



Future Improvements: To enhance the recommendation system, further machine learning needs to be applied to incorporate user feedback and rating. The dataset needs to be expanded, to include more diverse cuisine and more features can be created based on both: new data and user feedback. Lastly, refine the algorithms for better accuracy.

Summary: “Our project successfully demonstrates the potential of data-driven recipe recommendations. By leveraging user data and advanced analysis techniques, we can provide personalized suggestions that cater to individual tastes and preferences.”

Impact: This recommendation system can improve the user experience, making it easier for people to access and learn new recipes based on their unique preferences.