



Analysis on Indian Food

- Anoosha Shetty



Dataset details

- The dataset contains the data on 255 traditional and famous dishes of India.
- Dish attributes: Name, Ingredients, Diet, Preparation Time, Cooking Time, Flavor_Profile, Course, State, Region
- States covered are listed below:
 - **Southern States**: Andhra Pradesh, Karnataka, Telangana, Kerala, Tamil Nadu
 - **Northern States**: Punjab, Uttar Pradesh, Uttarakhand, Bihar, NCT of Delhi, Jammu & Kashmir, Haryana
 - **Eastern States**: West Bengal, Odisha
 - **Western States**: Rajasthan, Maharashtra, Gujarat, Goa
 - **Central States**: Chhattisgarh, Madhya Pradesh
 - **North Eastern States**: Assam, Tripura, Manipur, Nagaland

Data Analysis Steps

Data Extraction



Data downloaded from Kaggle

Data Cleansing



Data cleansing and transformation performed using Python. Found the top 10 ingredients used.

Data Visualization



Data Visualization on Power BI

Observations



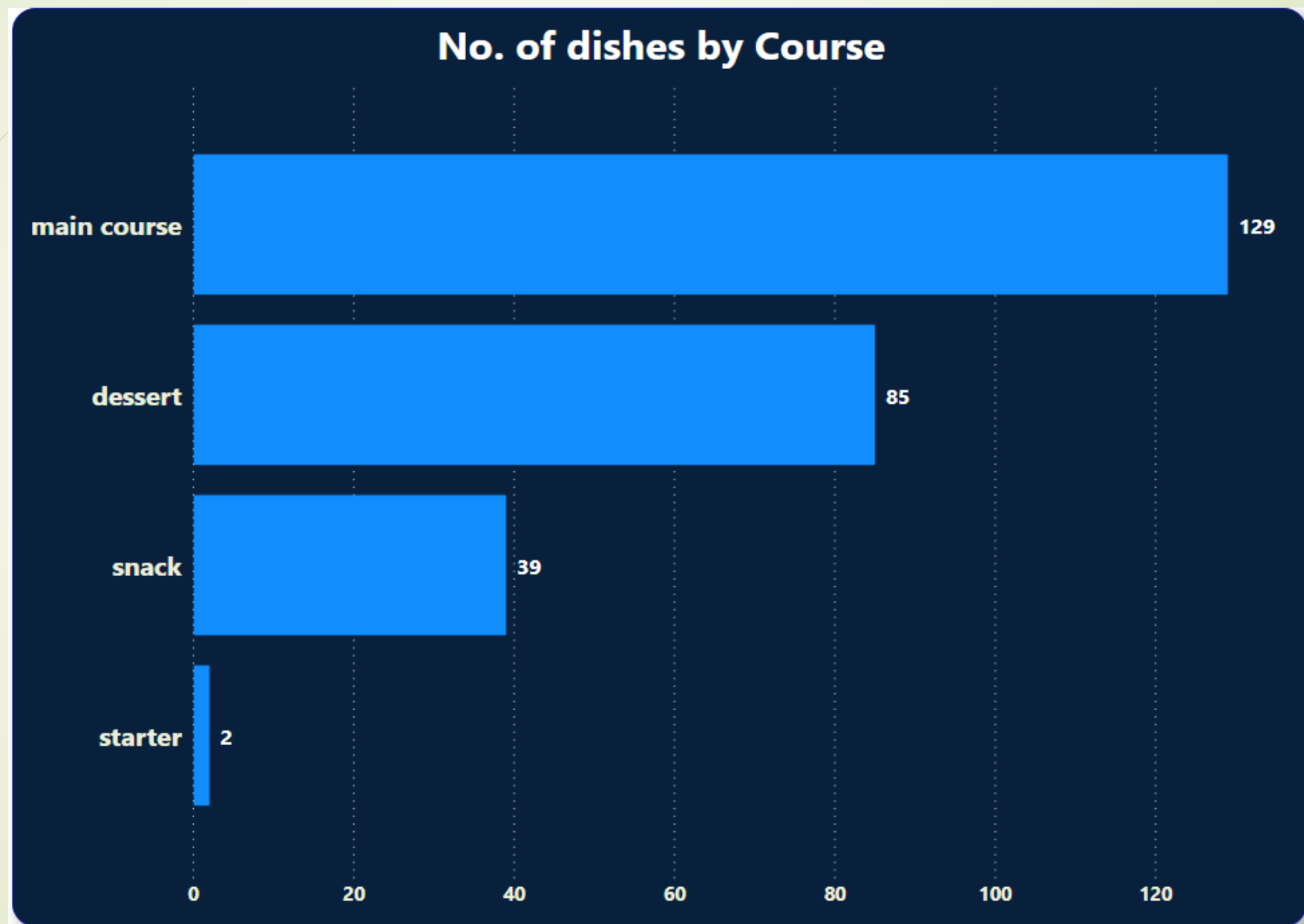
Findings evaluated and presented against each visualization in the subsequent slides

Top 10 Ingredients

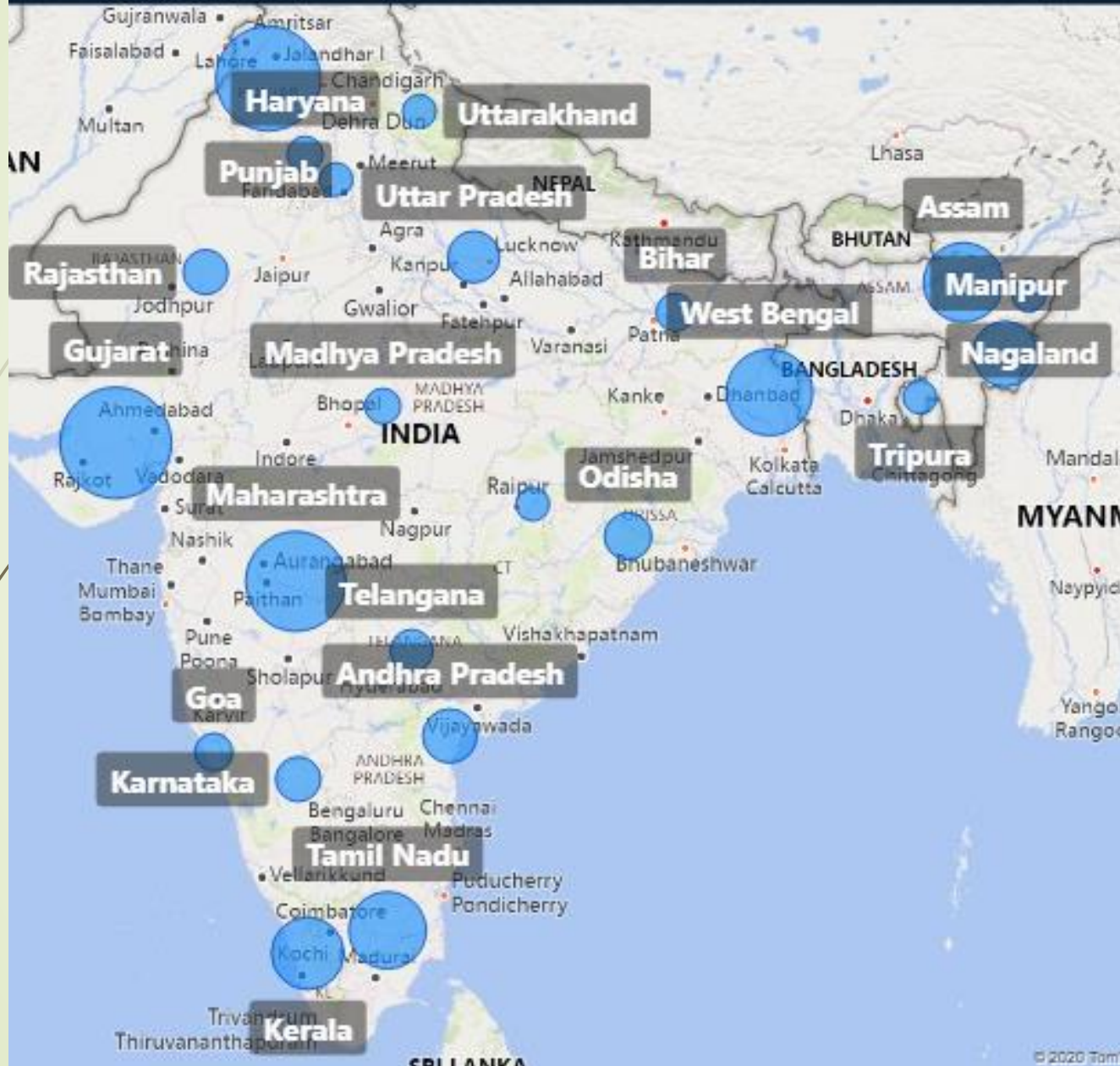
The top
10 ingredients used
along
with their frequency
of use across the
dishes

Ingredient Name	Frequency
Sugar	48
Garam Masala	40
Ghee	34
Ginger	29
Jaggery	27
Besan	27
Urad Dal	25
Curry Leaves	25
Tomatoes	24
Maida	24

Count of dishes by Course



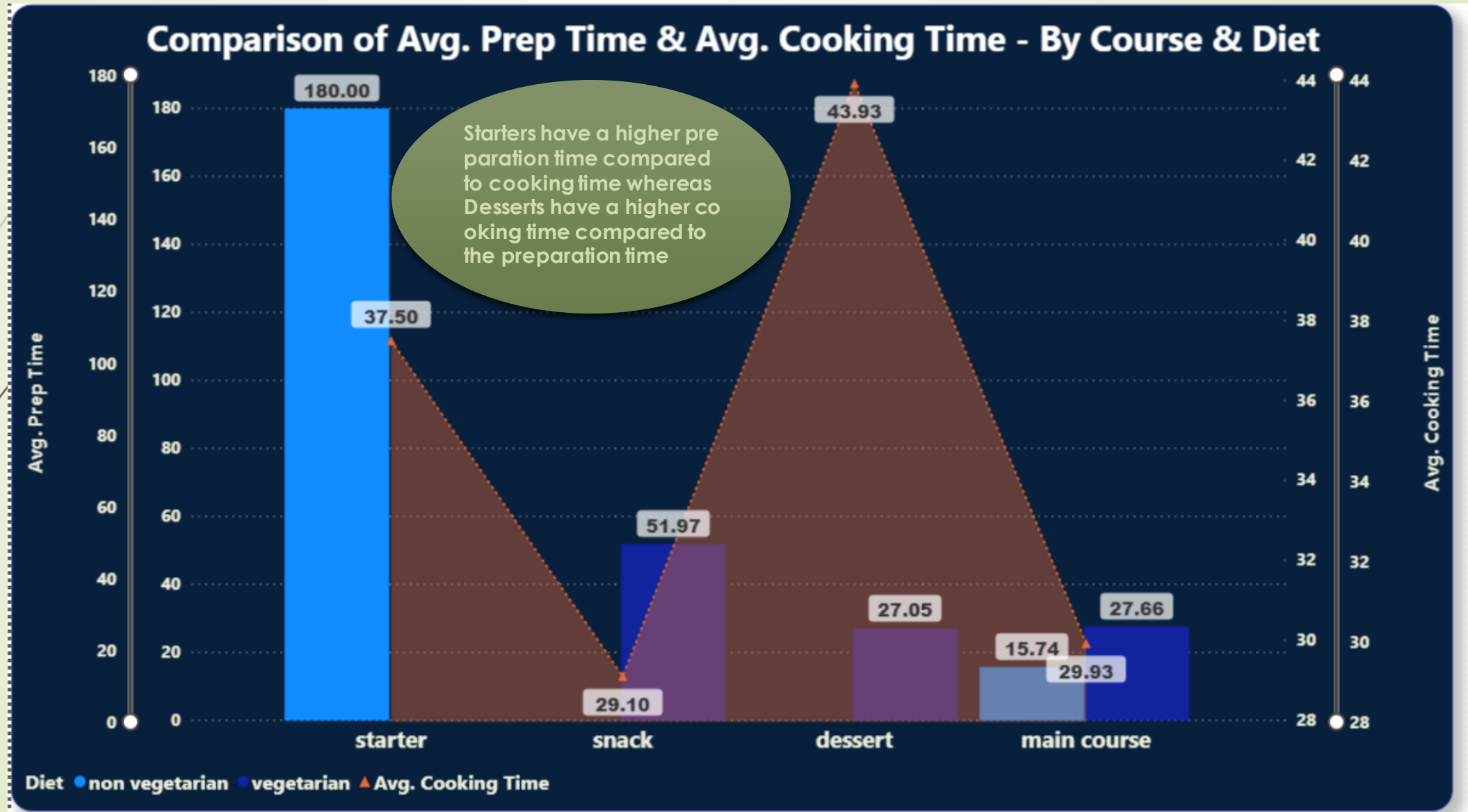
Country wise Distribution



Major contributors region wise

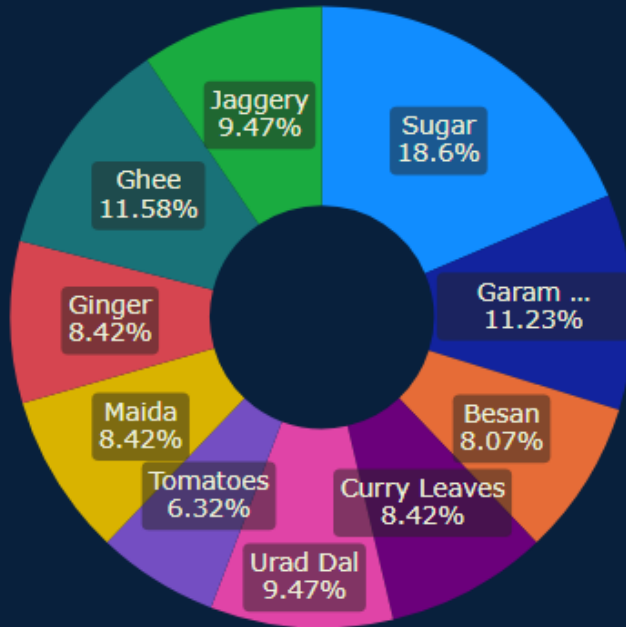
- North Region ==> Punjab
- West Region ==> Gujarat & Maharashtra
- South Region ==> Tamil Nadu & Kerala
- East Region ==> West Bengal
- North East Region ==> Assam
- Central Region ==> Madhya Pradesh

Preparation Time vs Cooking Time

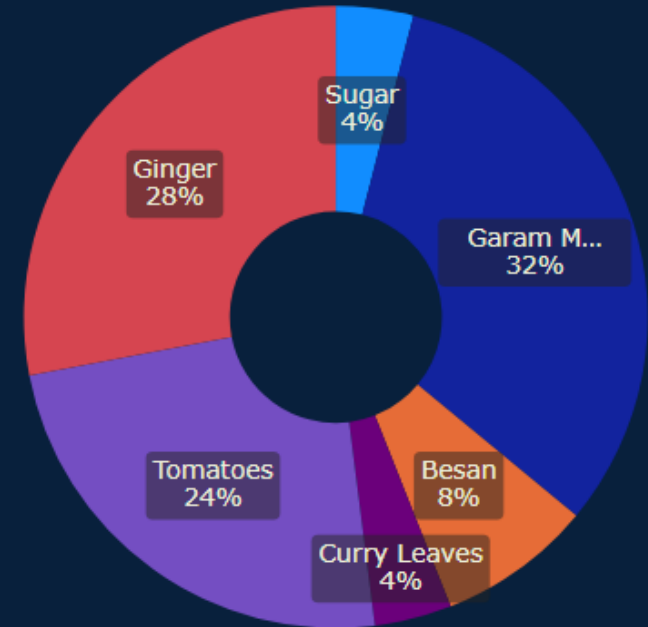


Ingredient Composition by Diet

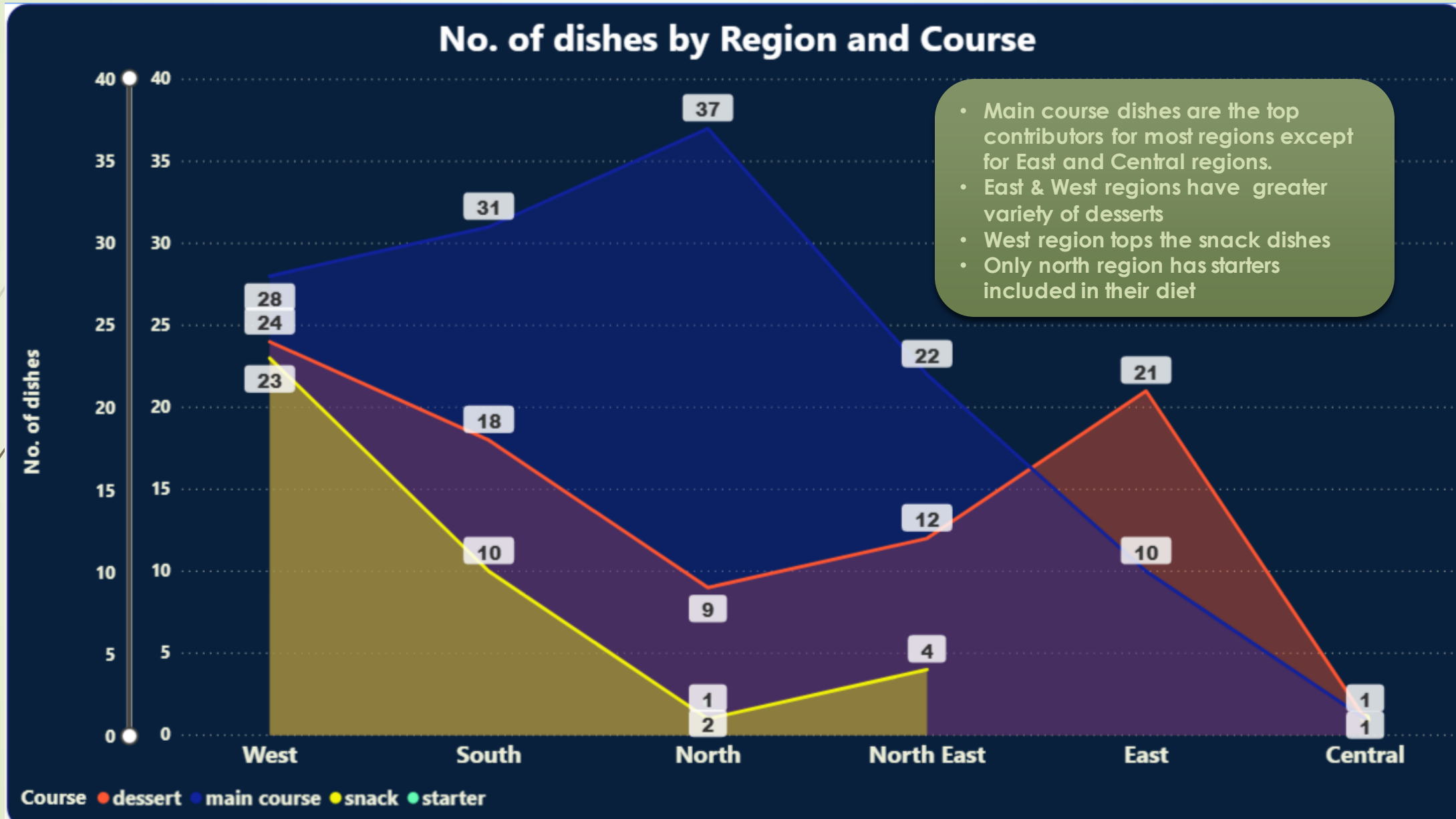
Vegetarian



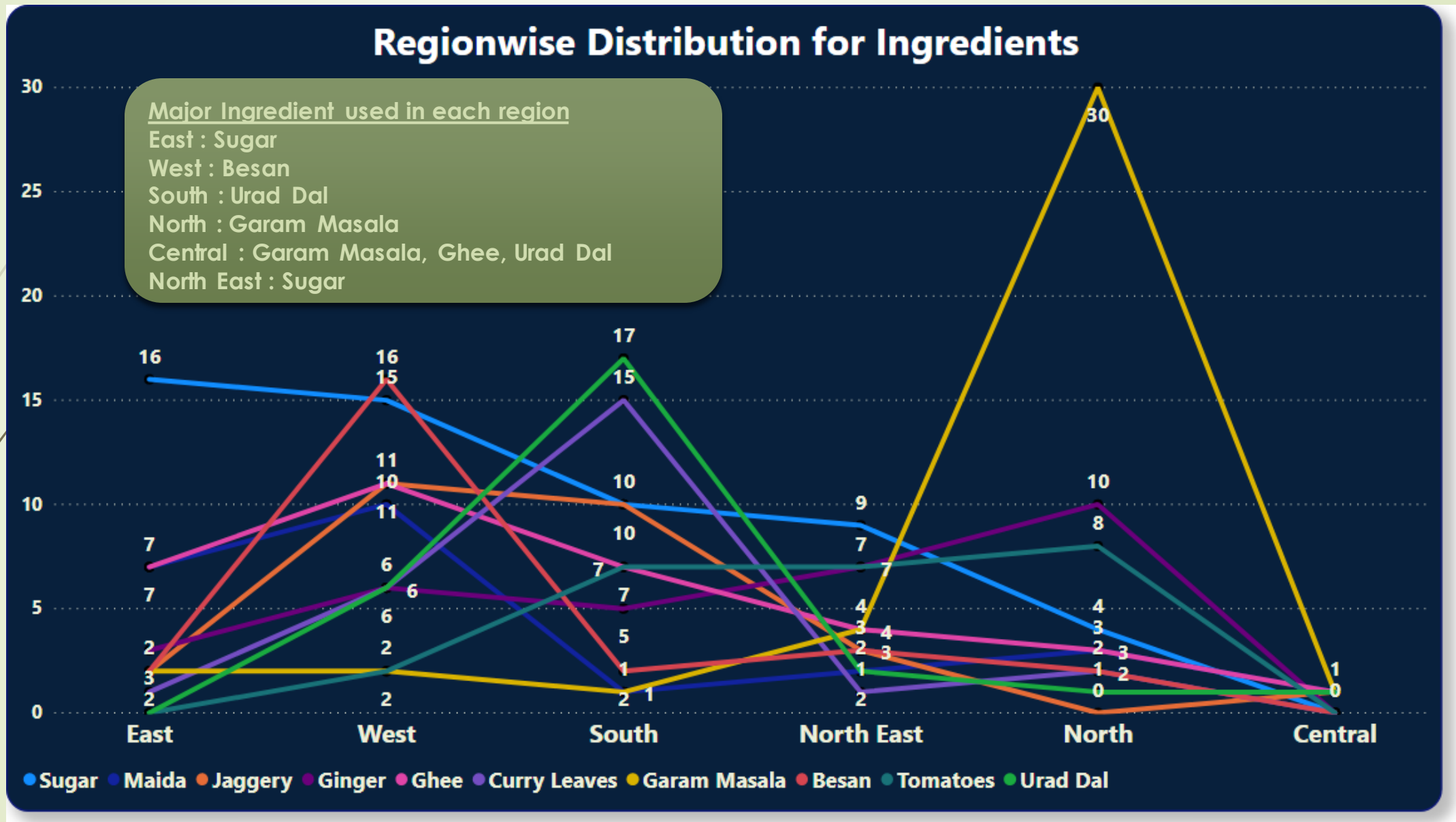
Non-Vegetarian



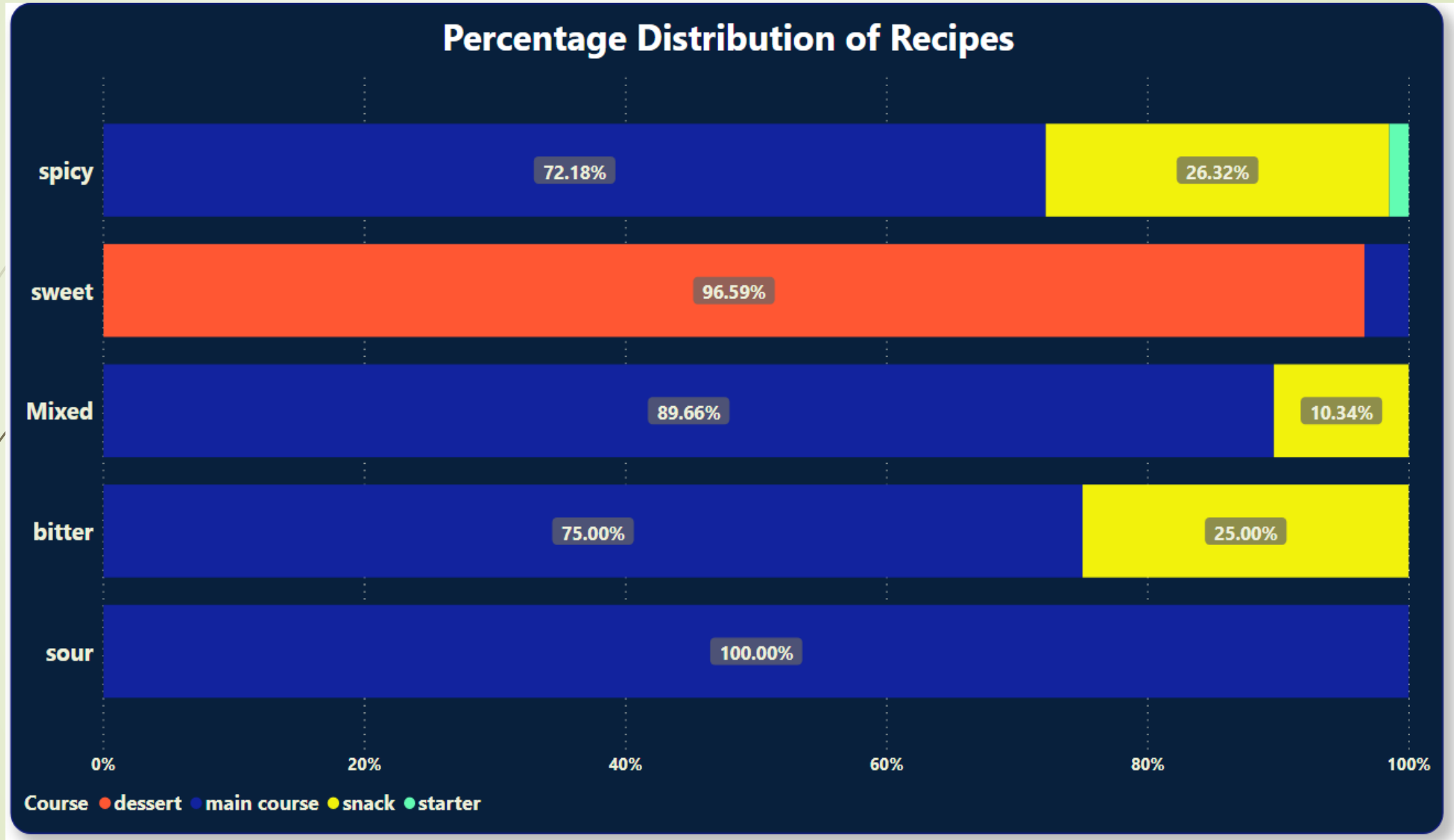
Region wise and Course wise Distribution



Top Ingredients Region wise Distribution



Percentage Distribution of Recipes



Conclusion & Future Analysis Scope



Main course forms the major contributor followed by Dessert, Snack & Starters. Hence we find their higher presence in almost all the slides. West region (mainly Gujarat & Maharashtra) are leading in region wise contribution



Implement Machine Learning algorithms to derive more insights