Data Collection and Preprocessing Phase

Date	6 October 2024
Team ID	SWTID1726832093
Project Title	Analysis of Amazon Cell Phone Reviews Using NLP Technique
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	This project aims to analyze Amazon cell phone reviews using Natural Language Processing (NLP) techniques to uncover customer sentiments, trends, and key themes influencing user satisfaction.
Data Collection Plan	Data will be collected from publicly available datasets on Kaggle and through scraping from the Amazon website (if necessary). The focus will be on reviews related to various cell phone models.

Raw Data Sources Identified	Web Scraped Reviews = **Review s** Raw reviews collected through web scraping from Amazon, focusing on the latest reviews of popular cell phone models.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissi
Amazon Customer Reviews	A dataset containing customer reviews for various products, including cell phones. It includes text	https://www.kaggle.com/datasets/karkavelrajaj/amazon-sales-dataset	CSV	1GB	Public
Kaggle Mobile Reviews	A collection of reviews specifically focused on mobile phones, with detailed user feedback and	https://www.kaggle.com/datasets?search=amazon+dataset	CSV	0.5GB	Public
…	…	…	…	…	…