|  |
| --- |
|  |
| **Team 2**  Abhishek Chintamani  Amit Shukla  Rupali Agarwal  Smaranika Sikdar  Varun Ananthula |

# Data Collection & Pre-Processing

# Group Assignment

# Domain : Indian Stock Market

# 

|  |
| --- |
|  |
|  |

Contents

[Data Collection & Pre-Processing 1](https://microsoft-my.sharepoint.com/personal/amshuk_microsoft_com/Documents/CPP%20Assignment%20Document.docx#_Toc103694951)

[Group Assignment 1](https://microsoft-my.sharepoint.com/personal/amshuk_microsoft_com/Documents/CPP%20Assignment%20Document.docx#_Toc103694952)

[Domain : Indian Stock Market 1](https://microsoft-my.sharepoint.com/personal/amshuk_microsoft_com/Documents/CPP%20Assignment%20Document.docx#_Toc103694953)

[Executive Summary 3](#_Toc103694954)

[Domain & Seed Sources 4](#_Toc103694955)

[Structured & Unstructured Sources 5](#_Toc103694956)

[Download / Crawl / Collect Data 6](#_Toc103694957)

[Data Conversion from Original Sources to Structured Data 7](#_Toc103694958)

[Data Cleaning & Pre-Processing 8](#_Toc103694959)

[Observations / Insights & Analytics 9](#_Toc103694960)

[Strategy to enhance data with Crowd Sourcing Methods 10](#_Toc103694961)

[References & Sources 11](#_Toc103694962)

## Executive Summary

###### Problem Statement

###### Proposed Solution

###### Challenges

## Domain & Seed Sources

###### Chosen Domain

###### Seed Sources

## Structured & Unstructured Sources

## Data Collection

###### Downloaded Data

###### Crawled Data

## Data Format Conversion

## Data Cleaning & Pre-Processing

## Observation & Insights

## Enhancing data with Crowd Sourcing Methods

## References & Sources