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

# Data Collection & Pre-Processing

# Group Assignment

# Domain : Indian Stock Market

# 

|  |
| --- |
|  |
|  |

Table of Contents

[Data Collection & Pre-Processing 1](file:///C:\Personal\Google%20Drive\Next%20Level\ISB\Term%201\Data%20Processing\Assignment\DCPP-Assignment\DCPP%20Assignment%20Strategy%20Document.docx#_Toc103723969)

[Group Assignment 1](file:///C:\Personal\Google%20Drive\Next%20Level\ISB\Term%201\Data%20Processing\Assignment\DCPP-Assignment\DCPP%20Assignment%20Strategy%20Document.docx#_Toc103723970)

[Domain : Indian Stock Market 1](file:///C:\Personal\Google%20Drive\Next%20Level\ISB\Term%201\Data%20Processing\Assignment\DCPP-Assignment\DCPP%20Assignment%20Strategy%20Document.docx#_Toc103723971)

[1. Executive Summary 3](#_Toc103723972)

[2. Domain & Seed Sources 4](#_Toc103723973)

[3. Structured & Unstructured Sources 5](#_Toc103723974)

[4. Data Collection 6](#_Toc103723975)

[5. Data Format Conversion 7](#_Toc103723976)

[6. Data Cleaning & Pre-Processing 8](#_Toc103723977)

[7. Observation & Insights 9](#_Toc103723978)

[8. Enhancing data with Crowd Sourcing Methods 10](#_Toc103723979)

[9. References & Sources 11](#_Toc103723980)

## 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