**Week 1 Project Documentation**

**Date:** 22-01-2025

**Title:** Exhaustive Analysis of Indian Agriculture- Week 1 Progress

**Objective:** The major goal of this project is to create an exhaustive analysis of Indian Agriculture using Power BI. In Week 1 tasks, I focused on the ETL process which refers to Extraction, Transformation, Loading process. It involved extracting, cleaning, transforming, and loading the given dataset into Power BI to get it ready for analysis and visualization.

**Dataset:**

The data set consists of about 19 years of Indian Agriculture data. It is not pure and contains empty and unnecessary columns. It is a CSV file and contains about 73827 rows and 12 columns out of which 4 columns are blank and 1 contains row names which are to be removed in the cleaning process. It contains the states and districts as well as the relevant data about year, season, crop, area and production

**Steps Taken in Week 1:**

**1] Installation and exploration of PowerBI:**

I installed the PowerBI application from the website into my device. Then I explored the various tabs and options it provided.

**2] Data Set Extraction and Analysis:**

The data set containing details of the various Indian states and their districts’ agriculture related statistics spanning across the season, year, type of crop, area and production. After downloading the dataset which was in CSV format, I used the get data option in the Home tab to extract data into the PowerBI application.

**3] Data Cleaning:**

Using Power Query Editor, I promoted the headers, checked for missing values using the view tab, removed unnecessary columns like blanks and row\_id by using the remove columns option that were not relevant to the analysis.

**4] Data Transformation:**

I standardized the column names using all lowercase and \_. I also ensured that the data types governing the columns were correctly identified.

**5] Data Loading into Power BI:**

Using the close and apply option in Power Query Editor, I loaded dataset into Power BI.

The Power BI file was then saved as “ExhaustiveAnalysisOnIndianFarmeAgriculture.pbix” for further processing.

**Output Screenshots:**

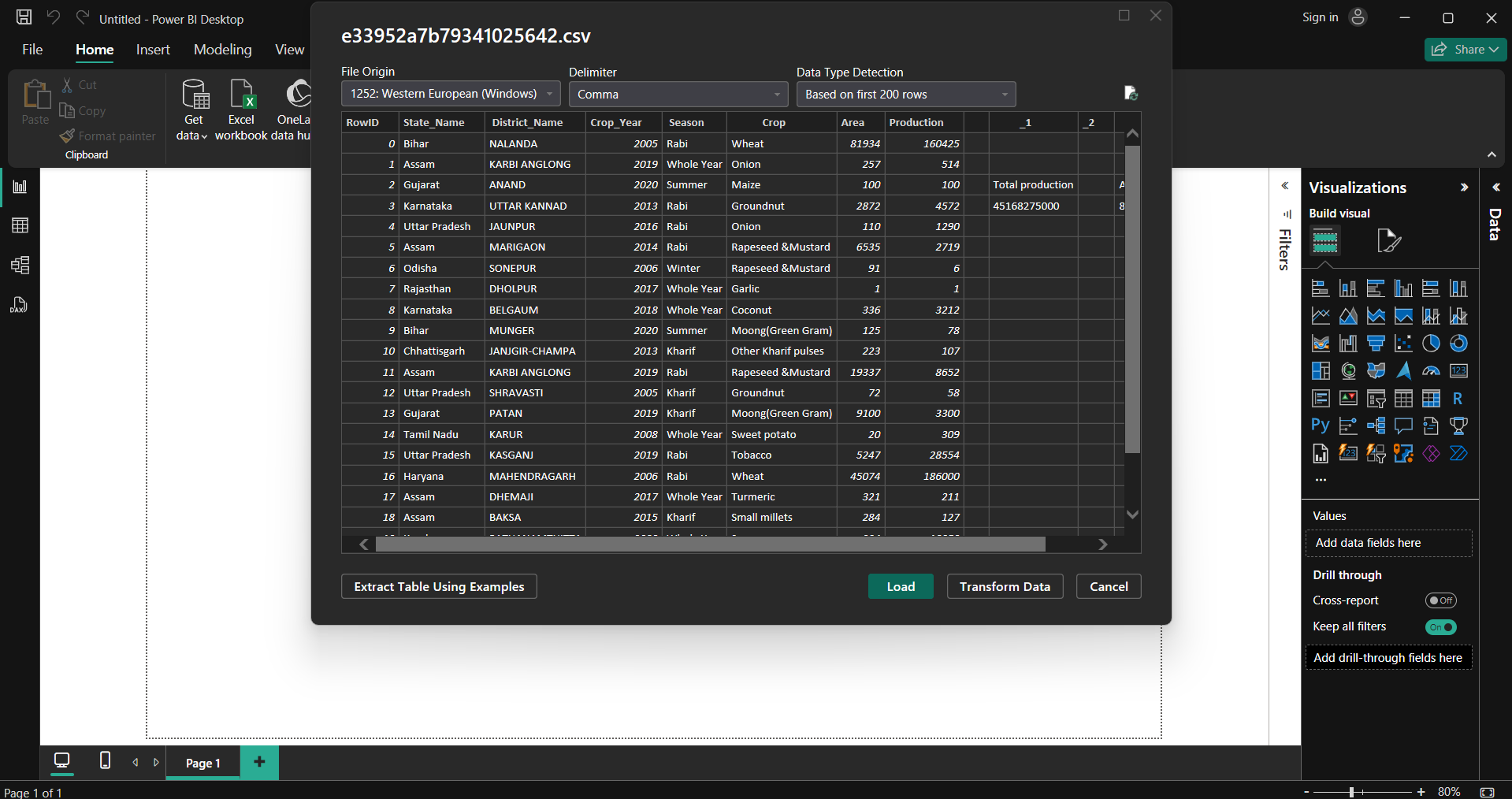


Fig- Extraction/ Importing of CSV file

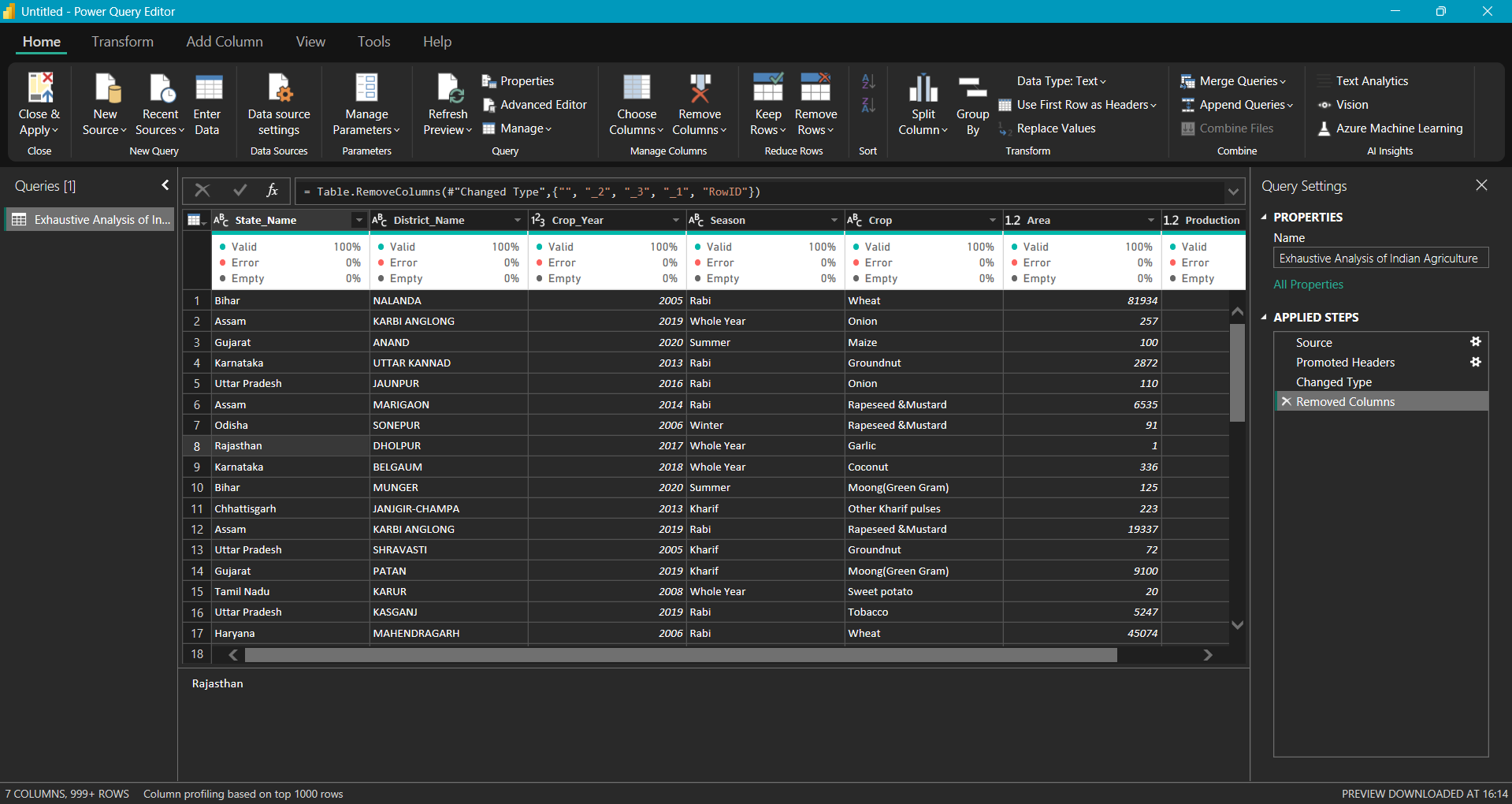


Fig- Analysis of Missing Values and removing of unnecessary and blank columns in Power Query editor

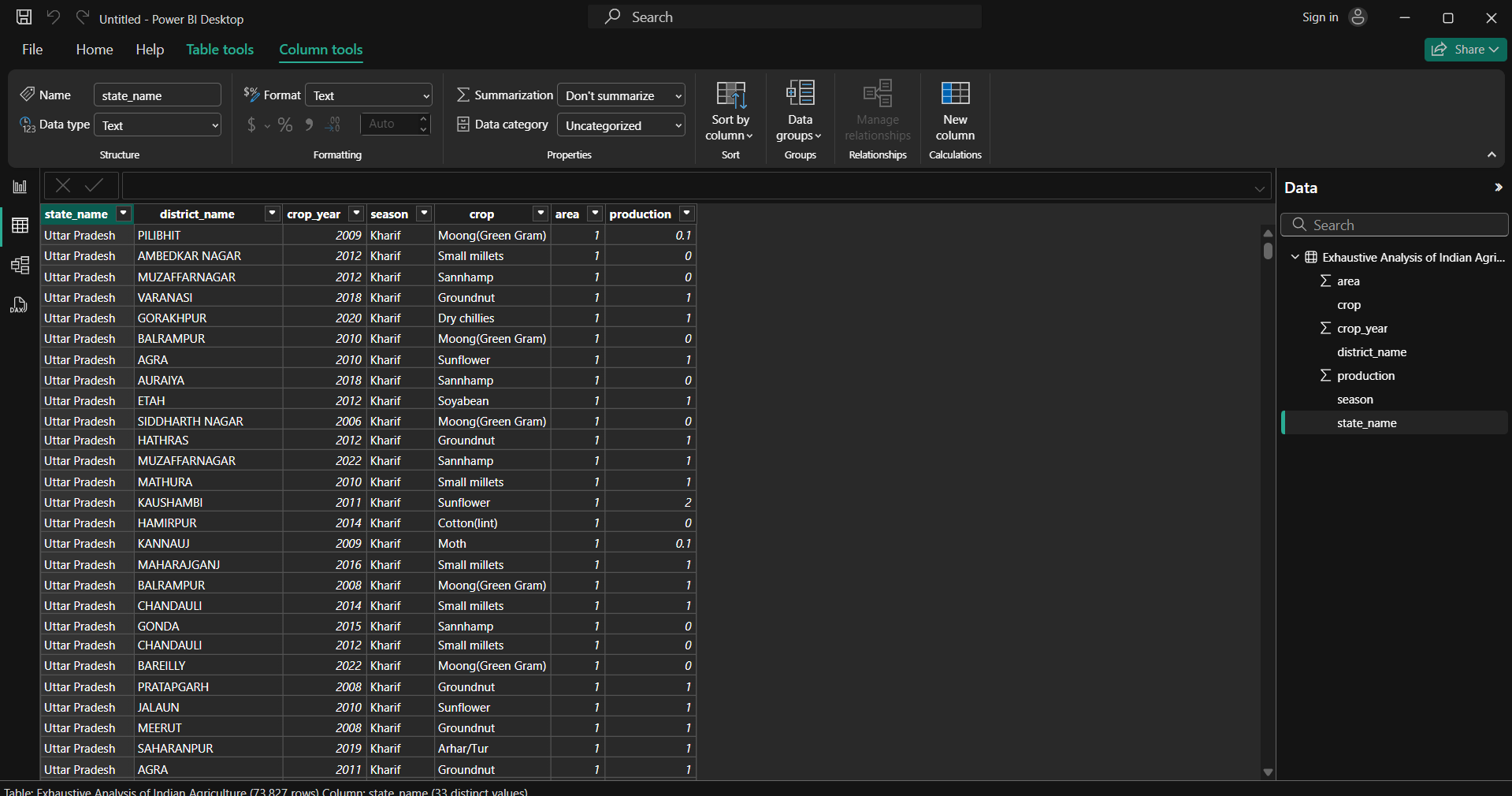


Fig-Loading of cleaned and Transformed data into PowerBI

**Conclusion:**

The ETL process was used to prepare the Dataset for analysis and creation of Dashboard. It provided pure data which is not ready to be used for visualization.