**Diamond Data Processing**

**Project Overview**

The Diamond Data Processing project is designed to automate the pre-processing and organization of diamond-related data. It takes raw data in the form of Excel files, extracts information from specific columns, and processes it to create structured datasets. The primary use case for this project is in the diamond industry, where data from various sources needs to be standardized for further analysis and reporting.

**Features**

- Data extraction from Excel files.

- Standardization of data values using a keyword map.

- Calculation of the final price for each diamond.

- Data export to separate Excel files for different purposes.

**Technical Details**

1. **Prerequisites**

Before getting started, ensure that you have the following software and libraries installed on your system:

- Python 3.x (https://www.python.org/downloads/)

- Pandas (https://pandas.pydata.org/pandas-docs/stable/getting\_started/install.html)

- xlwings (https://docs.xlwings.org/en/stable/installation.html)

1. **Installation**

Follow these steps to set up the project on your local machine:

1. Unzip the project into yours system
2. Dump all files to be processed in folder “unprocessed”
3. Install python (if you don’t have)
4. Install Pip (if you don’t have)
5. Install libraries (Pandas, xlwings, re, configparser, xlsxwriter etc)
6. Go to “src” folder. Update config.ini based on your system configuration.
7. Open terminal and navigate to folder: src
8. Run following command: - python main.py
9. Agent and vendor files are ready in their respective folders.
10. **Usage**

Follow these steps to carry out regular tasks on yours machine:

1. Dump all files to be processed in folder “unprocessed”
2. Open terminal and navigate to folder: src
3. Run following command: - python main.py
4. Agent and vendor files are ready in their respective folders.