# **NasugBUY: An E-commerce Platform for Local Businesses in Nasugbu with Cluster Analysis**

# **Data Integration Approach**

The process that we will use will be an Extract, Transform, Load (ETL) process. We are going to collect data from different sources, then we will transform them and apply calculations and analytics and will load the data so it can be usable. The table below shows the data sources that we will use and where do we plan to use them:

|  |  |
| --- | --- |
| **DATA SOURCE** | **USAGE** |
| Business Owners | The dataset from the business owners will be their information and the business information itself. |
| Nasugbu Locals | Same as business owners, it will be the information, but mostly their preferred products/services/business. |
| Nasugbu Tourism Industry and Services Office | We can’t have the Mamaraka dataset due to privacy, so the data that will come from them will be the idea on what will help the system when it comes to promoting business since they will be the Admin of our system. |
| Business Permit and Licensing Office – Steps for Securing Business | We were able to officially have a copy of steps in acquiring a business permit, since we are planning to attach the steps on our system to help the business owners make their business official and secured. |
| Department of Trade and Industry – Suggested Retail Price | We were able to officially have a copy of their SRP and it will be used to help the system when it comes to pricing. |

The table below shows the analytical tools that we may or may not use. (Still in discussion on which one to use.):

|  |  |
| --- | --- |
| **ANALYTICAL TOOLS** | **USAGE** |
| Tableau | We plan to use Tableau since Tableau uses the k-means algorithm for clustering, and it will be an advantage since we are using cluster analysis on our system. |
| R Studio | K-means and hierarchical clustering can also be applied on R Studio, so that’s why it is one of the list. |
| Jupyter Notebook | It is one of the tools that we may consider since we have also found out that k-means clustering can also be applied using Jupyter. |