**Duplication Typecasting**

**Problem statement:**

Q1. For the given dataset perform the type casting (convert the datatypes, ex. float to int)

Q2. Check for the duplicate values, and handle the duplicate values (ex. drop)

Q3. Do the data analysis (EDA)?

Such as histogram, boxplot, scatterplot etc

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| InvoiceNo | StockCode | Description | Quantity | InvoiceDate | UnitPrice | CustomerID | Country |
| 536365 | 85123A | WHITE HANGING HEART T-LIGHT HOLDER | 6 | 12/1/2010 8:26 | 2.55 | 17850 | United Kingdom |
| 536365 | 71053 | WHITE METAL LANTERN | 6 | 12/1/2010 8:26 | 3.39 | 17850 | United Kingdom |
| 536365 | 84406B | CREAM CUPID HEARTS COAT HANGER | 8 | 12/1/2010 8:26 | 2.75 | 17850 | United Kingdom |
| 536365 | 84029G | KNITTED UNION FLAG HOT WATER BOTTLE | 6 | 12/1/2010 8:26 | 3.39 | 17850 | United Kingdom |
| 536365 | 84029E | RED WOOLLY HOTTIE WHITE HEART. | 6 | 12/1/2010 8:26 | 3.39 | 17850 | United Kingdom |
| 536365 | 22752 | SET 7 BABUSHKA NESTING BOXES | 2 | 12/1/2010 8:26 | 7.65 | 17850 | United Kingdom |
| 536365 | 21730 | GLASS STAR FROSTED T-LIGHT HOLDER | 6 | 12/1/2010 8:26 | 4.25 | 17850 | United Kingdom |
| 536366 | 22633 | HAND WARMER UNION JACK | 6 | 12/1/2010 8:28 | 1.85 | 17850 | United Kingdom |
| 536366 | 22632 | HAND WARMER RED POLKA DOT | 6 | 12/1/2010 8:28 | 1.85 | 17850 | United Kingdom |

**Ans:-**

**Data types: -**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of features** | **Description** | **Type** | **Relevance** |
| InvoiceNo | InvoiceNo | Nominal | Not Relevant |
| StockCode | StockCode | Nominal | Relevant |
| Description | Description | Nominal | Relevant |
| Quantity | Quantity | Count | Relevant |
| InvoiceDate | InvoiceDate | Count | Relevant |
| UnitPrice | UnitPrice | Ratio | Relevant |
| CustomerID | CustomerID | Nominal | Not Relevant |
| Country | Country | Nominal | Relevant |