# **System Integration Testing (SIT) & User Acceptance Testing (UAT) Documentation**

Group Number – **08**

Project - **Medium Data Analytics Leaderboard System**

Team Members – **Piyush Kumar Sultania**

**Gauri Narendra Bhale**

**Nameera Mohsin**

## SIT Testing\_1: Data Transformation Test

In Original Dataset (Before Cleansing)

|  |  |
| --- | --- |
| **Title Name** | **10 Best Android Emulators for PC and Mac of 2017** |
| From Source dataset | Tag\_android = 1 |
| Publication | NULL |



In Cleaned Dataset (Post Cleansing)

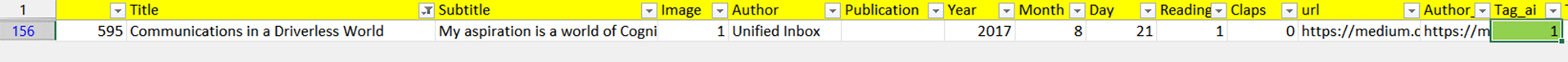
|  |  |  |
| --- | --- | --- |
|  |  | **Activity Performed** |
| From Transformed dataset | Category = "android" |  |
| Publication | No Publication | Pivot the tags |



## SIT Testing\_2: Data Transformation Test

In Original Dataset (Before Cleansing)

|  |  |
| --- | --- |
| **Title Name** | **Communications in a Driverless World** |
| From Source dataset | Tag\_ai = 1 |
| Publication | NULL |



In Cleaned Dataset (Post Cleansing)

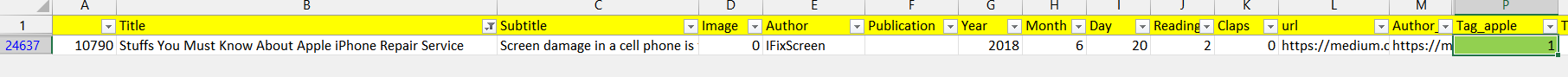
|  |  |  |
| --- | --- | --- |
|  |  | **Activity Performed** |
| From Transformed dataset | Category = "ai" |  |
| Publication | No Publication | Pivot the tags |



## SIT Testing\_3: Data Transformation Test

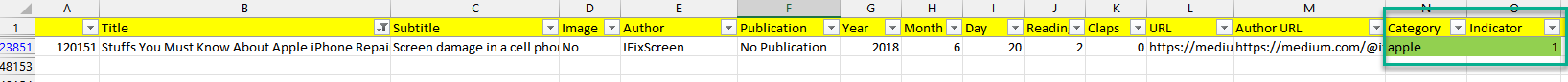
In Original Dataset (Before Cleansing)

|  |  |
| --- | --- |
| **Title Name** | **Stuffs You Must Know About Apple iPhone Repair Service** |
| From Source dataset | Tag\_apple = 1 |
| Publication | NULL |



In Cleaned Dataset (Post Cleansing)

|  |  |  |
| --- | --- | --- |
|  |  | **Activity Performed** |
| From Transformed dataset | Category = "apple" |  |
| Publication | No Publication | Pivot the tags |



## SIT Testing\_4: NULL Check

*Note: This test case is prepared on data set of 50000 records!*

|  |
| --- |
| **Before data wrangling** |
| Total NaN Title Entries: 1849 |
| Total NaN Author Entries: 0 |

Graphical user interface, application

Description automatically generated

|  |
| --- |
| **After data wrangling** |
| Total Null Title Entries: 0 |
| Total Null Author Entries: 0 |

Graphical user interface, text, email

Description automatically generated

## SIT Testing\_5: Count Check!

|  |  |
| --- | --- |
| **Record count (Before Cleaning)** | **500000** |

Graphical user interface, application

Description automatically generated

Text

Description automatically generated

Text

Description automatically generated

|  |  |  |
| --- | --- | --- |
| **Record count (After Cleaning)** | **482068** |  |
| 1 | Total dropped NaN Title Entries | 17932 |
| 2 | Total dropped NaN Author Entries | 0 |

Graphical user interface, text, application

Description automatically generated

## SIT Testing\_6: Capitalize and Titling Fields:

**Before data wrangling**

Publication:MILLENNIALSTIMES // NEW ORDERA picture containing table

Description automatically generated

**After data wrangling**Graphical user interface, application

Description automatically generated

## SIT Testing\_7: Converting Image field into Boolean (Yes/No):

**Before data wrangling**

|  |  |
| --- | --- |
| **Title Name** | **Online Animation: Mixamo vs Norah AI** |

A picture containing calendar

Description automatically generated

**After data wrangling**

Graphical user interface, application

Description automatically generated

## SIT Testing\_8: Renaming Fields into business fields:

**Before data wrangling**

|  |  |
| --- | --- |
| **Field Name** | **Unnamed: 0** |

A picture containing table

Description automatically generated

**After data wrangling**

Graphical user interface, application

Description automatically generated

## SIT Testing\_9: Missing Value Handling (Empty/Blanks and NULL) :

**Before data wrangling**

|  |  |
| --- | --- |
| **Field Name** | **Publication** |

A picture containing table

Description automatically generated

Graphical user interface, application

Description automatically generated**After data wrangling**

Graphical user interface, application

Description automatically generated

## SIT Testing\_10: Deriving Fields from source columns:

**Before data wrangling**

|  |  |
| --- | --- |
| **Field Name** | **Tag\_\* (90 columns)** |

A picture containing table

Description automatically generated

**After data wrangling**

|  |  |
| --- | --- |
| **Field Name** | **Category** |

* *We pivot the Tag\_\* columns and created a derived field called “Category” and picked all the records which have story-tag = 1 only!*
* *Also, we cleaned the fields by trimming the Tag\_ keyword from each column value and capitalize them before loading into target file.*

*Ex: Tag\_ai (column) 🡪 Ai (Row value)*

Graphical user interface, application

Description automatically generated

## SIT Testing\_12: CQL vs Excel (Data Integration Testing)

|  |  |
| --- | --- |
| **Field Name** | **Field Value** |
| Category |  |
| Count (From .csv) - Unique | 71 |
| Count (From Neo4j post Load) - Unique | 71 |

**From Source File (.csv)**

Table

Description automatically generated

**From Neo4j GraphDB**

A picture containing text

Description automatically generated