**Data migration**

Move the data from excel to a new application

**Scenario**:

New students get registered to the university; the data gets populated in spreadsheets; that contains the student information broadly in 4 categories: personal details, course details, logistics info and contact details.  
university has decided to move all the data to a new ERP system as it will be one main system where everyone can interact; ERP on college management system reduces the most of the human work that are done earlier to managing the college system.

This movement of data is a repetitive process for every college resource and requires a lot of human work and is a time consuming process.

**Solution:**

RPA (robotic process automation)  
We can deploy virtual workers who mimic human workers. Robot's productivity is 100% without any errors and bot can handle multiple tasks very well

**RPA Stages:**

1: Understand the data

-data standards  
 -data health  
 -how will the data be used tomorrow   
 - what is missing  
 -de-duplicate  
 -transform

Validation:

-relevance: Is it relevant to its intended purpose?  
 -accuracy: Is it correct  
 -integrity: Does it have a coherent, logical structure?  
 -consistency: Is it consistent and easily to understand?  
 -validity: Is it within acceptable parameters for the business?  
 -accessibility: Can it be easily accessed and exported to the target application?

why validations:

-improved corporate performance  
 -increased competitive advantage  
 -reduced uncertainty & risk  
 -reduced delays & waiting time  
 -measurable & accurate view of data  
 -better accountability

2: Extraction of data  
 -move data to a new spreadsheet

3: transformation  
 -match the data to its new forms; and ensure that metadata reflects the data in each fields

4: Cleansing  
 -de-duplicate;   
 -run tests  
 -address any corrupted data

5: validation   
 -test and re-test that moving data to the target location provides the expected response

6: Loading the data into the new system   
 -transfer the data into the new system and review for errors again

**Types of data:**

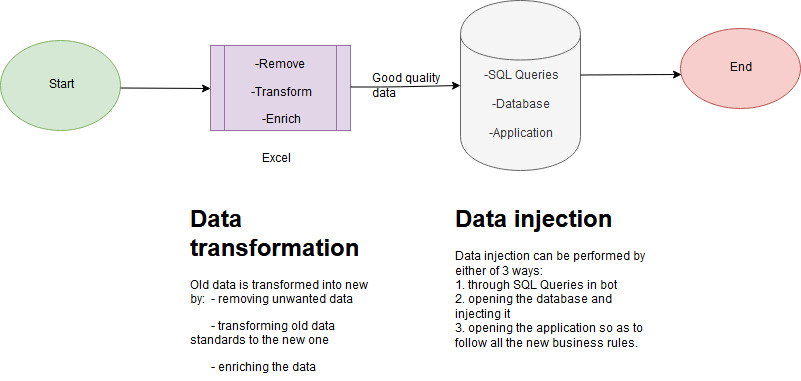
1: The data that’ll be removed:  
 Few data from the source need to be removed as it might be redundant or useless;  
 eg: Age.

2: The data that’ll be moved as is:  
 Personal information need not require any touch up  
 eg: personal info, phone number  
 3: Data to be transformed:  
 few data requires some touch up   
 eg: gender from ‘M|F’ to ‘Male|Female’, extracting zip code from address; data format from DDMMYY to YYYYMMDD

4: Enrich:  
 some data will be missing from the original doc and needs to be populated ;  
 eg: missing course details can be populated like course duration and the elective subjects

**Loading of Data:** The cleansed, good quality data will be injected to the new application by either of the 3 ways:

1: Bot executing the SQL queries;  
 2: Bot opening the database and injecting the data;  
 3: Bot logging on the new application and filling up the information (all the new business rules will be followed without even looking at them.)



**Phases**:

1: Bot fetches the data from spreadsheet performs macros cleanses the data then injects to the new application.

2: Bot fetches the data from multiple sources and transforms it then injects.

3: Bot is now able to convert the data from an unstructured format (fetching course details from a PDF or the university website) to formatted one.