QUOTE EXTRACTION AUTOMATION API

APR’ 22 – JUNE’ 22

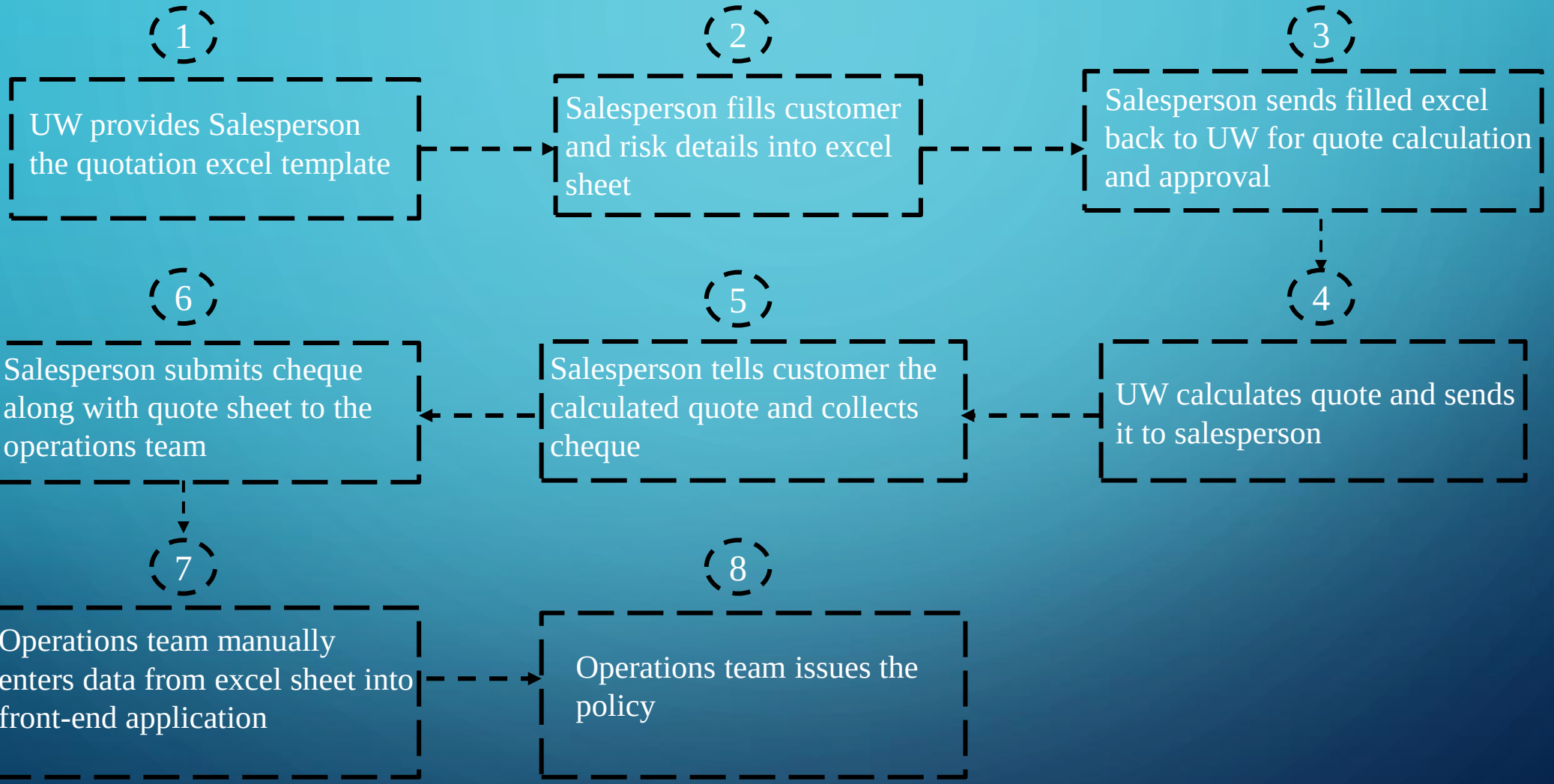
AGENDA

- Objective
- Current Process
- Challenges with Current Process
- Modified Process
- Advantages of Modified Process
- Managing Redundancy in Lead ID
- Products
- API Code Flow
- Demo
- Software Specifications
- Acknowledgements

OBJECTIVE

Create an API to automate the process of validation, extraction, and storage of data from Insurance Quote Excel into Database, thereby removing manual data entry and expediting the policy issuance process.

CURRENT PROCESS

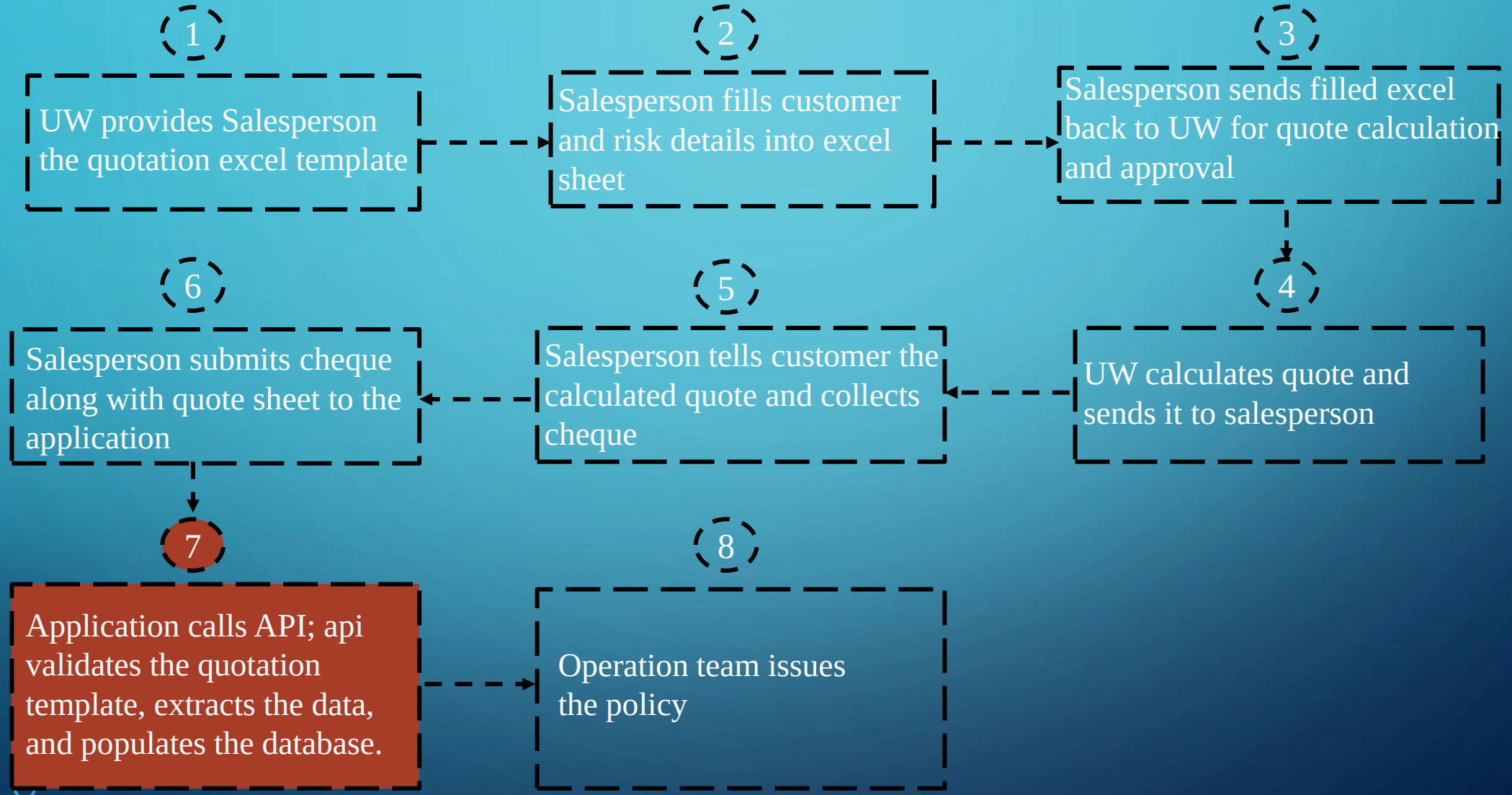


CHALLENGES WITH CURRENT PROCESS

- Excel quote template inconsistency:
  - Salesperson might have changed some formulas in the excel sheet.
  - Salesperson might be using the obsolete template.

- Time consuming: Manual data entry from excel into application for policy issuance.
- Manual Data Entry is error prone.

MODIFIED PROCESS



ADVANTAGES OF MODIFIED PROCESS

- 4 Factor Validation:
  - File is in **.xlsx** format
  - Checks for **unique hidden encrypted value** in a random cell in excel file
  - Checks for template format using **5 constant labels** in the right cells in excel sheet e.g. hazard category, natcat score, graded retention, mailing address, and occupancy.
  - Checks for the **creator and title** of the sheet from sheet properties.
- Automated Data Entry from Excel to Database:
  - Saves time:
    - ❖Current Process – **At least 2 hours** to generate 1 proposal
    - ❖Modified Process – **At most 2 minutes** to generate 1 proposal
  - No Data entry errors

MANAGING REDUNDANCY IN LEAD ID

- Whenever a customer reaches out to the insurance company for buying a product, a unique lead is created for that customer, which includes his/her basic details.
- It could happen that the same customer uploads the same excel file twice or thrice to out api, leading to multiple entries of records with the same lead id.
- So how to deal with this redundancy problem?
- We have created an “isactive” column in each table in the database. Whenever data is uploaded into the database, the api first checks if that lead id already exists. If it does, the “isactive” column of all previous records with that lead is made 0, and the new incoming records get the lead id 1. In this way, the company knows that the customer’s most recent activity are the records that have the lead id as 1.

PRODUCTS

1. Standard Fire & Special Perils Policy (SFSP)
2. Burglary Insurance
3. Plate Glass Insurance

For each Product the following sections were read:

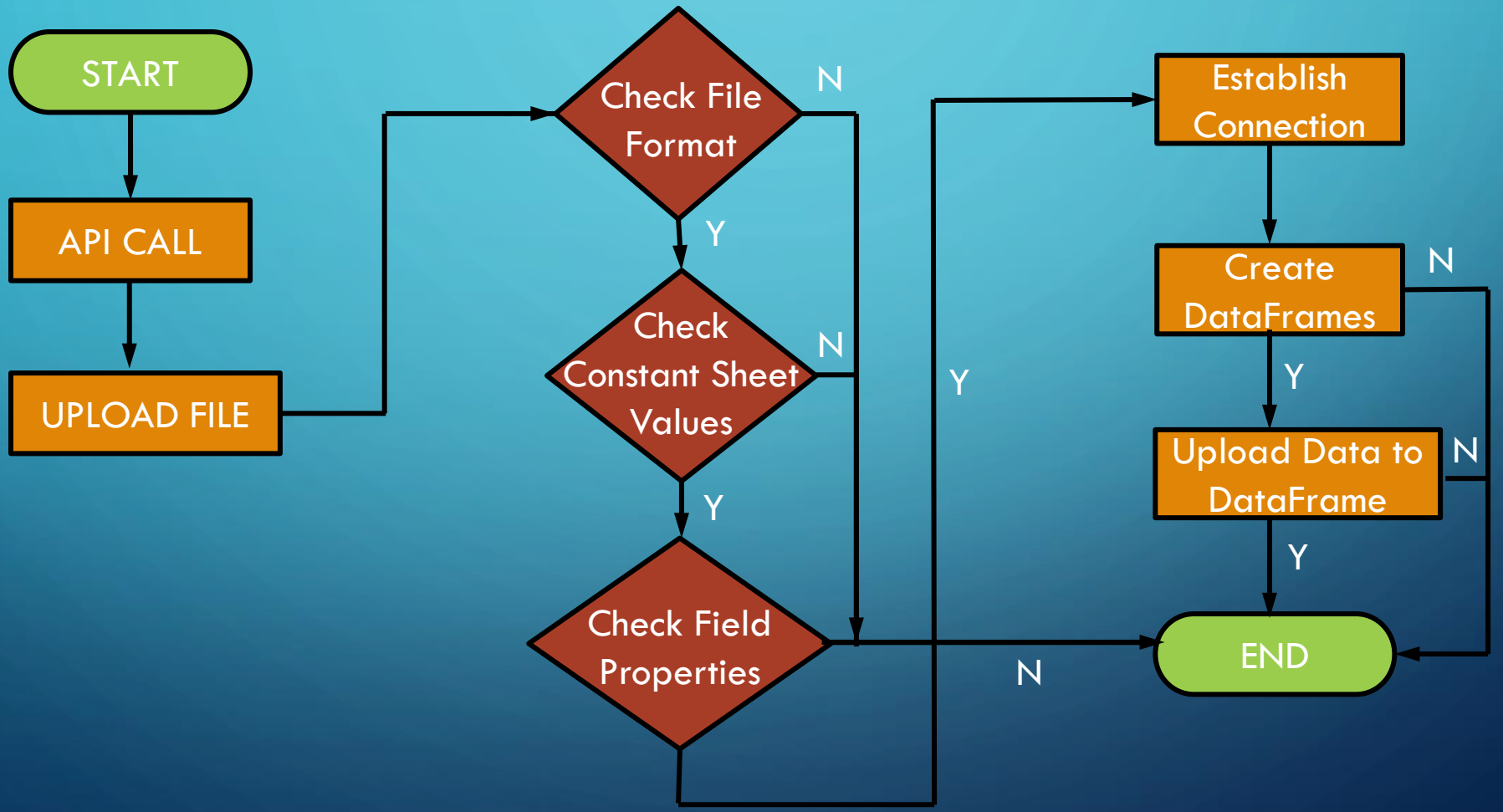
- Customer Details
- Main Policies
- Add-On Covers
- Terms & Conditions
- Deductibles, Warranties, Clauses & Conditions
- Exclusions & Subjectivities
- Supplementary Clauses & Conditions

API CODE FLOW

FILE UPLOAD

QUOTE FORMAT VALIDATION

QUOTE EXTRACTION

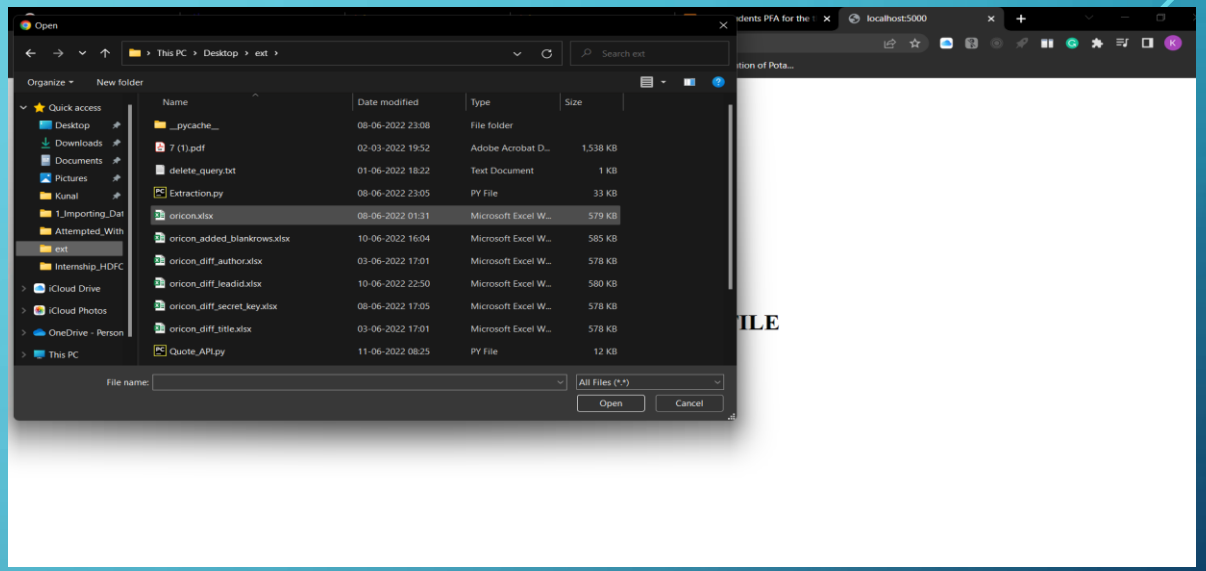


DEMO

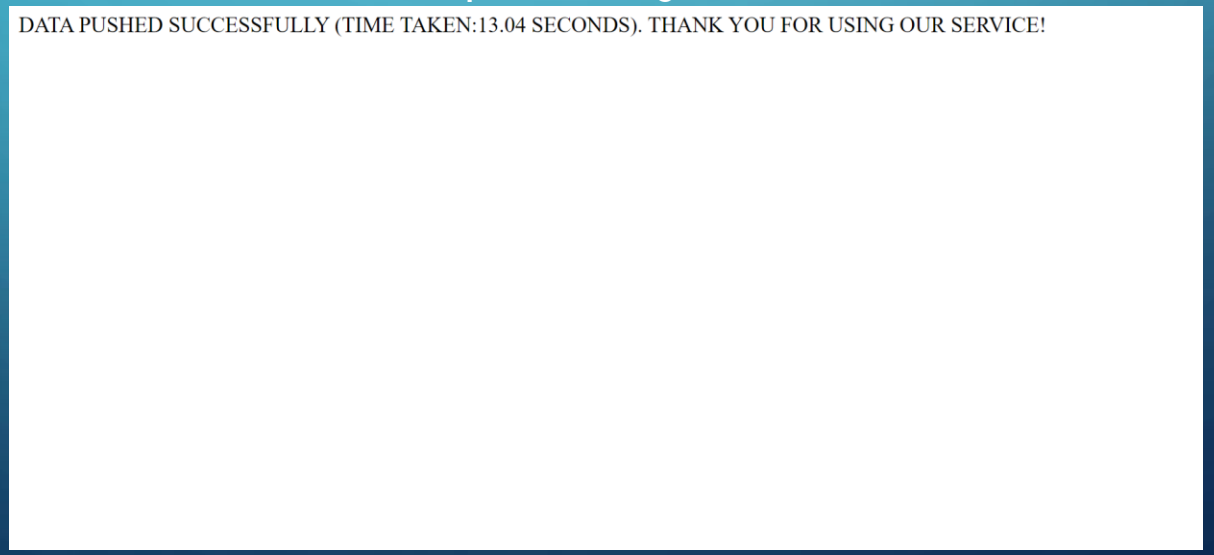
DEMO

Front-end to upload quote excel file

Select the quote file & submit



Output message



SOFTWARE SPECIFICATIONS

- Python – 3.8.8 with Oracle – 19c as Database
- Flask – 2.1.2
- Pandas – 1.4.2
- NumPy – 1.22.4
- Os – 0.6.3
- Sqlalchemy – 1.3.24 (Using this version is critical as some of the functions for executing queries using python from this version are deprecated in higher versions)
- Openpyxl – 3.0.9
- Werkzeug – 2.1.2
- Cx_Oracle – 8.3.0

ACKNOWLEDGEMENTS

We would like to express our gratitude to **Senior VP-IT Mr. Naresh Jha & HDFC ERGO**, who gave us this golden opportunity to do this wonderful project of Quote Extraction Automation API. We learned many new concepts in python and SQL programming, which have added to our skillset and have helped us grow. Secondly, we would also like to thank our mentors **Mr. Ketan Khandagale** and **Mr. Swapnil Gunjal** for assisting us in completing this project and providing help and support at every step of the project. Lastly, we would also like to thank our **parents** for helping us finalize this project within the stipulated time frame. We are grateful for getting this opportunity. Thank you.

THANK YOU