

REST APIs for DB services using Django

A thesis submitted to the department of

Computer Science & Engineering

of

International Institute of Information Technology, Bhubaneswar

in partial fulfilment of the requirements

for the degree of

Bachelor of Technology

by

Bhushan Prakash Khanale

(B516025)

under the supervision of

Harshad Mulmuley & Prof. Tapan Kumar Sahoo



Department of Computer Science and Technology

International Institute of Information Technology Bhubaneswar

Bhubaneswar Odisha - 751003, India

Undertaking

I declare that the work presented in this thesis titled REST APIs for DB services using Django, submitted to the Department of Computer Science and Engineering, International Institute of Information Technology, Bhubaneswar, for the award of the Bachelors of Technology degree in the Computer Science and Engineering, is my original work. I have not plagiarised or submitted the same work for the award of any other degree. In case this undertaking is found incorrect, I accept that my degree may be unconditionally withdrawn.

Bhushan Prakash Khanale

B516025

Certificate

This is to certify that the work in the thesis entitled *REST APIs for DB services using Django* by *Bhushan Prakash Khanale* is a record of an original research work carried out by her under my supervision and guidance in partial fulfillment of the requirements for the award of the degree of *Bachelor of Technology* in *Computer Science & Engineering*. Neither this thesis nor any part of it has been submitted for any degree or academic award elsewhere.

Harshad Mulmuley
Senior Software Engineer
Turtlemint Pvt. Ltd.

Tapan Kumar Sahoo
Associate Professor, Computer Science
IIIT, Bhubaneswar

Acknowledgement

The elation and gratification of this seminar will be incomplete without mentioning all the people who helped me to make it possible, whose gratitude and encouragement were invaluable to me. I would like to thank God, almighty, our supreme guide, for bestowing his blessings upon me in my entire endeavor. I express my sincere gratitude to Harshad Mulmuley & Prof. Tapan Kumar Sahoo for his guidance and support and students of my class for their support and suggestions.

Bhushan Khanale

B516025, CE

Abstract

Most of the databases now are shared between different tenants giving it a more complex architecture. Hence any updates being made to the database have to be properly authenticated and verified that the changes are for that specific tenant only. This project report introduces the process of creating a REST API service to manage database changes with integrated authentication using Django. REST is acronym for **RE**presentational **S**tate **T**ransfer. It is architectural style for distributed hypermedia systems. Django is an open-source high-level Python Web framework that encourages rapid development and clean, pragmatic design. By the features of Django and Django REST Framework these updates to the database are much simpler and protected. ***Keywords:*** rest, django, database

Contents

Abstract	v
1 Introduction	2
1.1 Background	2
1.2 Significance	3
1.3 Method used	3
1.4 Limitations	3
1.5 Project Structure	4
2 Preliminary	5
2.1 Django	5
2.1.1 Models	5
2.1.2 Views	6
2.1.3 Templates	6
2.2 Postgresql	7
2.3 Social Auth	7
3 Project Architecture	9
3.1 An overview of the project architecture	9
3.1.1 Database tables	9
3.1.2 CRUD operations	10

3.1.3	User permissions	10
3.2	Business Logic	12
4	Development Process	13
4.1	Initial setup	13
4.2	Create models	14
4.3	Create serializers	15
4.4	Add CRUD operations	16
4.5	Add views	16
4.6	API testing	17
4.6.1	Unittests	18
4.6.2	Postman tests	18
5	Turtlemint Projects with Django	20
5.1	Turtlemint Assistant Project	20
5.1.1	The primary objective	21
5.1.2	The features	22
5.2	MintAcademy	25

Chapter 1

Introduction

In any service dealing with the database it becomes extremely important to have a constant database structure in place before moving on towards the business logic. In Django we define the service in terms of **app**, **models**, **views** and **services**. These four parts represent the core logic service. **Views** take care of the exchange of the request and response objects from APIs. Usually when a API is called, a request object is sent to the server containing information about the request being made. The server then has to return the appropriate Response object which then the browser parses and outputs for the user. This exchange between request and response is a part of Views.

1.1 Background

Turtlemint has a separate database which records most of the things related to insurance policy issuance. This data is very volatile and is expected to change every month. Due to this, it becomes harder to change the database everytime there is a change in the information. To handle this issue, the purpose of the project is to create a new service which would wrap the infor-

mation change in terms of database calls and let the user seamlessly update the information.

1.2 Significance

The new service will be able to handle all information changes related to the database. Moreover the service would have an integrated authentication and authorization which allows multiple users to use this service at a time. Previously, someone from the development team had to intervene with the data team to manually create database queries and update accordingly. This process was not only time consuming but also was inefficient. The new service would solve this issue and would allow the data team itself to update the database.

1.3 Method used

The service is built using Django and Django Rest Framework (DRF) which are two Python packages built for faster development of database-driven web applications. Django is also open-source and allows users to modify the report, modify any bugs if they found any. This helps for long term support applications. Django has three major parts: **models**, **views** and **templates**. Models are used to create database schema, views contain the business logic and templates are used for user interface.

1.4 Limitations

Django being open-source does help in most issues. Although, since Django was built to reduce the development time significantly it might still not have

all features of a system with independent database architecture. Django also introduces the concept of migrations which are a set of database schema changes maintained as a set of files. These migrations can be difficult to manage if an applications is prone to lot of database changes.

1.5 Project Structure

Django has already defined its project structure. Every Django project has some applications. Every applications represents set of logic related to one purpose or business objective. Every project can have any number of applications inside it. There is a common `settings.py` file which is used for managing settings for all applications. The basic structure of the project can be represented as below:

```
project
├── settings.py
├── app1
│   ├── models
│   ├── views
│   └── templates
└── app2
    ├── models
    ├── views
    └── templates
```

Chapter 2

Preliminary

There are three important parts of this project:

- Django (for API developement)
- Postgresql (for database requirements)
- Social Auth (for authentication)

2.1 Django

Django takes care of integrating the database, authentication and authorization of the application. For this, Django asks us to construct each of models, views and templates for the application. Its important to note that all of the three parts are to be defined for every application in the project.

2.1.1 Models

Every model represents a table in the database. We will be forming up the basic model structure and follow them throughout the entire project. Every

model can be defined in the `models.py` file which is recognized by django or in the `models` module. The overall model structure would look like this:

```
project
├── app1
│   └── models
│       ├── modela.py
│       ├── modelb.py
│       └── modelc.py
└── app2
    └── models
        ├── modela.py
        └── modelb.py
```

2.1.2 Views

Every view represents the exchange of data between every API call. Every view gets a request object and is expected to return a response object. The request contains all information regarding the API call including the user who initiated it. Also all of the required parameters for the API call are passed in the same request object.

Views too can be distributed across multiple files for convenience.

2.1.3 Templates

Templates are used for the user interface which are automatically rendered by Django in html. By default, Django comes with the `admin` template. The admin page on Django lets user to modify any of the existing models and settings. Multiple templates can be added representing the user interface. The templates are also convenient in authentication since we can use variables inside templates.

2.2 Postgresql

We've used Postgresql for our database requirements. Because postgres can scale on high requirement environment and is one of the best options for production databases. The postgresql can be installed as a separate package and is available on all platforms including Linux, Windows and MacOS making it ideal for deployment and development purpose.

Postgres related settings can be configured inside `settings.py` file inside project directory. We can also specify the requirements as a part of the environment file to avoid exposing the credentials. The database setting is generic in Django meaning you can use the database identically as any other database driver. Hence, it is very easy to switch between database if we ever wanted to.

2.3 Social Auth

For authentication there is a separate package known as *Python Social Auth* which takes care of our social authentication. Social authentication is important because many users now try to avoid remembering username and password and hence, using one of the social auth method will bypass the need to remember usernames and passwords. This authentication is based on one of the social media platforms including Google, Facebook, Twitter, etc. These platforms provide authentication flow known as *OAuth*. The common user fields like email address, name are automatically obtained from these platforms and a *token* is created for the user to communicate with the application.

Everyone in Turtlemint has their email with domain *turtlemint.com* which lets us use Google OAuth2 to authenticate the user before logging

it on our application. The email address and name is obtained from Google and appropriate permissions are issued to the user based on his access level.

Chapter 3

Project Architecture

3.1 An overview of the project architecture

The project objective is to ease the process of updating and managing the database with proper authorization for the users. This adds up to few main components of the project which are:

- Database tables
- CRUD operations
- User permissions

3.1.1 Database tables

Currently in Turtlemint the entire database is in nosql MongoDB. Hence, the first challenge is to convert these nosql tables into relational postgres tables and import all data from the previous database to current database. The number of such tables is huge and hence some scripts are generated to copy the previous database tables into the current ones. We will take a look at these scripts later in the models section.

3.1.2 CRUD operations

Every table or model in a database has some basic CRUD (Create, Read, Update, Delete) operations associated with it. These are declared in the models class itself to avoid rewriting them again in the business logic. These operations handle the basic interaction with the database and can be used anywhere in the project scope.

```
from django.db import models

def create(**kwargs):
    turtlemint_model.objects.create(**kwargs)

def read(**kwargs):
    turtlemint_model.objects.get(**kwargs)

def delete(**kwargs):
    turtlemint_model.objects.get(**kwargs).delete()

class turtlemint_model(models.Model):
    """
    Model representing turtlemint meta information.
    """
    key = models.CharField(max_length=100, unique=True, blank=True)
    name = models.CharField(max_length=100)
    description = models.TextField(max_length=100)
    ...
```

3.1.3 User permissions

There are usually two ways we can create users based on their permission levels. One of the ways is to use Django's inbuilt `groups` which can create

separate groups for users and every group would then have a dedicated level of permission which would then be applied for every user in that group. This process is handy for most of the simple applications which don't have to deal with a lot of different user sets.

The other way is to define a `CustomUser` model overriding Django's inbuilt `User` model. This gives us much more freedom over the user characteristics and can then be used with different permission classes. As mentioned above, Django have some of the permission classes already defined while other can be manually created.

```
class TurtlemintUser(AbstractBaseUser):
    """
    This class represents Custom User Model,
    that overrides django builtin User model
    """
    key = models.CharField(max_length=100, unique=True, blank=True)
    email = models.EmailField(
        verbose_name='email address',
        max_length=255,
        unique=True,
    )
    first_name = models.CharField(max_length=100, blank=True)
    last_name = models.CharField(max_length=100, blank=True)
    password = models.CharField(max_length=200, blank=True)
    ...
```

Every user has some sort of permissions associated with it. These permissions are used to determine whether a user can interact with a certain database table or not. Django in itself provides `PermissionClasses` which can be used to determine the scope for the user. We can also create custom

permissions classes by extending the `BasePermission` class and overriding appropriate class methods.

An example of a `CustomPermission` class can be given as below:

```
from rest_framework.permissions import BasePermission
class IsSuperAdmin(BasePermission):
    """
    Allows access only to super admin users.
    """
    def has_permission(self, request, view):
        return bool(request.user and request.user.is_staff)
```

3.2 Business Logic

Most of the architecture is covered in the above sections, there are still some organization specific requirements which are to be fulfilled. These are all considered under business logic. The business logic just represents the actual code written to complete an objective or requirement imposed by the organization. At Turtlemint, we've number of requirements for which we need to split the project into several modules.

The most of the business logic should go in the `services/` directory. This is not a Django specific module, but is actually user defined. Hence, the name can be changed to whichever suitable. The business logic should be as isolated as possible since then it will be easier to write unittests against it. Unittests are written to check for every possible scenario and to check if its response is the one which is expected. Having a good code coverage will help maintainers in maintaining the software in long run.

Chapter 4

Development Process

4.1 Initial setup

The initial project setup consists of various steps including setting up the local development environment, Python and its dependencies and finally the initial Django project template. Django project can be initialized using the command `django-admin startproject mysite`. The command will create a new directory named `mysite` in your current directory. It represents the Django project. Once the project has been initialized, we need to create a new application within the project by the command `python manage.py startapp appname`. This will create a new directory inside your project directory representing a Django application.

Once the application has been initialized, we can move towards the actual development of the application. This process involves several steps including creating models, views, serializers, services etc. But before that it is important to setup the project settings correctly. You can select the database driver of your choice and modify the `settings.py` file in your project directory.

4.2 Create models

As mentioned previously before, models represent the database tables. These models can be defined in separate files or can be defined in a single file named `models.py` as required. Before starting to write the actual implementation it is critical to design and write the database schema first. Django provides you with an inbuilt `Model` class which you can extend to write your own models. Most of the methods in the `Model` class are reusable and hence we don't have to write custom methods for every class.

```
# models/turtlemint_model.py
```

```
class TurtleMintModel(models.Model):
    """
    Model representing TurtleMint meta model.
    """
    key = models.CharField(max_length=100, unique=True, blank=True)
    name = models.CharField(max_length=100)
    description = models.TextField(max_length=100)
    logo = models.CharField(max_length=200, null=True, blank=True)
    home_url = models.CharField(max_length=200, null=True)
    support_email = models.EmailField(blank=True, null=True)
    ...
```

```
# models/turtlemint_insurer_model.py
```

```
class InsurerModel(models.Model):
    """
    Model representing insurer meta model.
    """
    key = models.CharField(max_length=100, unique=True, blank=True)
```

```
name = models.CharField(max_length=100)
associated_vertical = models.CharField(max_length=200, null=True)
logo = models.CharField(max_length=200, null=True, blank=True)
home_url = models.CharField(max_length=200, null=True, blank=True)
support_email = models.EmailField(blank=True, null=True)
...
```

To make reflect these models into our database we need to create Django migrations. These migrations can be created using the command `python manage.py makemigrations`. This would create a new directory inside you app representing the database queries. Internally, Django parses the changes made to the models into their equivalent database queries. These created migrations are then ran using the command `python manage.py migrate` to reflect them in your database.

4.3 Create serializers

Serializers allow complex data such as querysets and model instances to be converted to native Python datatypes that can then be easily rendered into JSON, XML or other content types. Serializers also provide deserialization, allowing parsed data to be converted back into complex types, after first validating the incoming data.

Every models should have its own serializer, representing the construction of the JSON object through the model fields. Serializers are extremely useful to return the JSON object instantly without having to construct it manually. Django allows us to define the fields we want the serializer to construct the JSON with. Also, we can have multiple serializers for a model if we want a separate response based on the request.

An example of serializer can be given as below:

```
class TurtlemintModelSerializer(serializers.ModelSerializer):  
    """  
    Serializer for the TurtlemintModel.  
    """  
    class Meta:  
        model = TurtlemintModel  
        fields = ['id', 'key', 'name', 'description', 'language'...]
```

4.4 Add CRUD operations

Every model should have its own each of CRUD methods. These methods will be commonly shared and used by the views and business logic. Also these functions should be unittestable to make sure that they are doing exactly what is intended of them. Hence it is critical that these methods should be an early part of the development process.

4.5 Add views

Views take care of accepting the request and returning the appropriate response. These are defined inside the `views` module and can be split into various files and classes. Views can be written in two ways, having a separate class to handle all requests including GET, POST, DELETE, etc for a single endpoint or to have individual view methods for each type of request. Django supports both ways equally and hence any method should suffice based on the development need.

Examples of views:

```
class TurtlemintView(APIView):
```

```
"""
Turtlemint view.
"""

def get(self, request):
    return Response(Turtlemint.get_data(request.data),
                    status_code=status.HTTP_200_OK)

def post(self, request):
    response = Turtlemint.update(request.data)
    if response['success']:
        return Response(response, status_code=HTTP_200_OK)
    return Response(response, status_code=HTTP_400_BAD_REQUEST)
```

To make this operational for an endpoint, we need to add an entry in `urls.py` file for the same.

```
# urls.py
urlpatterns = [
    ...
    path('api/v1/turtlemint/meta', TurtlemintView.as_view()),
    ...
]
```

The views also link to the business logic for that API, this will be covered in the next chapter as a part of Turtlemint and its requirements.

4.6 API testing

The APIs can be manually tested using any framework like Postman. Django also provides a nice UI to be able to send requests. At turtlemint we've a culture of using Postman extensively for API testing and we've followed

the same. Although, apart from the Postman tests it is also necessary to construct unittests and view tests using Django's `TestCase` class.

4.6.1 Unittests

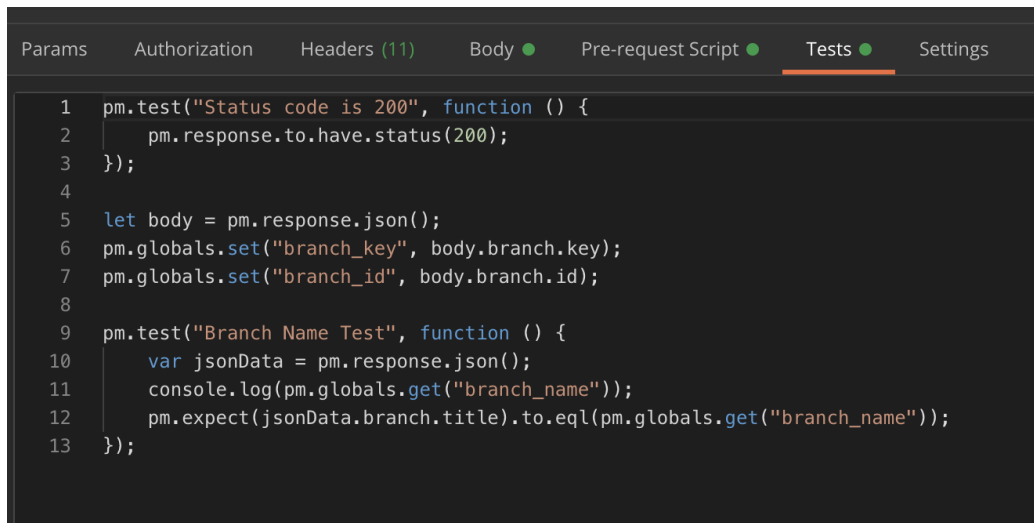
Django provides an extended version of `unittest.TestCase` class which automatically creates a test database and destroys it at the end of test execution cycle. Also every test is ran as a separate transaction hence allowing the database to rollback automatically once some error is encountered in the tests.

As per the default behaviour of unittest, every test file should start with `test` prefix and should correctly represent the file under testing. Tests can be written as a part of a class which may include `setUp` and `tearDown` function to represent to initiate, and clear some process at the start and the end of the test execution.

```
class TestTurtlemintModel(TestCase):
    def setUp(self):
        # Run something before running the testcase
        self.model = TurtlemintModel
    def test_get_records_count(self):
        # Get the output from the method under test
        result = get_records_count(self.model)
        self.assertEqual(result, 1) # 1 representing expected result
```

4.6.2 Postman tests

Unittests are always preferred, but for some reason if we need some sort of integration tests then postman tests can be written. Postman is an application used for API collection management. It introduces a lot of features

The image shows a screenshot of the Postman application's 'Tests' tab. The interface has a dark theme. At the top, there are several tabs: 'Params', 'Authorization', 'Headers (11)', 'Body', 'Pre-request Script', 'Tests' (which is selected and highlighted with an orange underline), and 'Settings'. Below the tabs, the test code is displayed in a monospaced font. The code consists of two test functions. The first function, 'Status code is 200', checks the response status. The second function, 'Branch Name Test', parses the response JSON, sets global variables for 'branch_key' and 'branch_id', logs the 'branch_name', and uses 'pm.expect' to verify that the 'title' of the branch matches the value stored in the 'branch_name' global variable.

```
1 pm.test("Status code is 200", function () {  
2   pm.response.to.have.status(200);  
3 });  
4  
5 let body = pm.response.json();  
6 pm.globals.set("branch_key", body.branch.key);  
7 pm.globals.set("branch_id", body.branch.id);  
8  
9 pm.test("Branch Name Test", function () {  
10   var jsonData = pm.response.json();  
11   console.log(pm.globals.get("branch_name"));  
12   pm.expect(jsonData.branch.title).to.eql(pm.globals.get("branch_name"));  
13 });
```

Figure 4.1: Sample Postman testcase

including one to write tests for the API collections. For every request you can mock the expected response and check if the response is correct or not. Although this might be helpful in some cases, we initially wrote some tests here but later due to introduction of unittests we dropped this process.

Chapter 5

Turtlemint Projects with Django

Majority of the Turtlemint codebase is written in Java. Hence, introduction of Python and Django in specific was very new to the team. As a part of the project I was working on two of their major projects namely, Turtlemint Assistant and MintAcademy. Both of these applications are deployed on production and are completely written in Django.

5.1 Turtlemint Assistant Project

The Turtlemint Assistant Project is one of the newest projects Turtlemint as developed. Me along with other interns are one of the core backend developers for this project. This project was written from scratch and involved all development processes discussed in last chapter.

The development processed involved majorly involved following activities:

- Understanding the existing database

- Replicating the existing database in relational model
- Starting a new Django project
- Developing the project to solve all requirements
- Take feedback from managers and project leads
- Testing the application
- Preparing unittests and integration tests
- Deploying application on staging server and then on production
- Solving issues, releasing bug fixes
- Feedback from users and adapting to their requirements

5.1.1 The primary objective

The primary objective of the project was to let the data team to conveniently update the insurer data in database. Since Turtlemint is an insurance company and acts as a broker between the insurance providers and customers, it has to record all insurer data in order to punch insurance policies on their behalf. This data is updated very frequently by the insurers and hence Turtlemint too has to adapt to those changes quickly to avoid punching wrong policies.

This was then acted upon us in order to provide them a common platform to handle database updates with a sophisticated yet simple user interface and reliable and stable system.

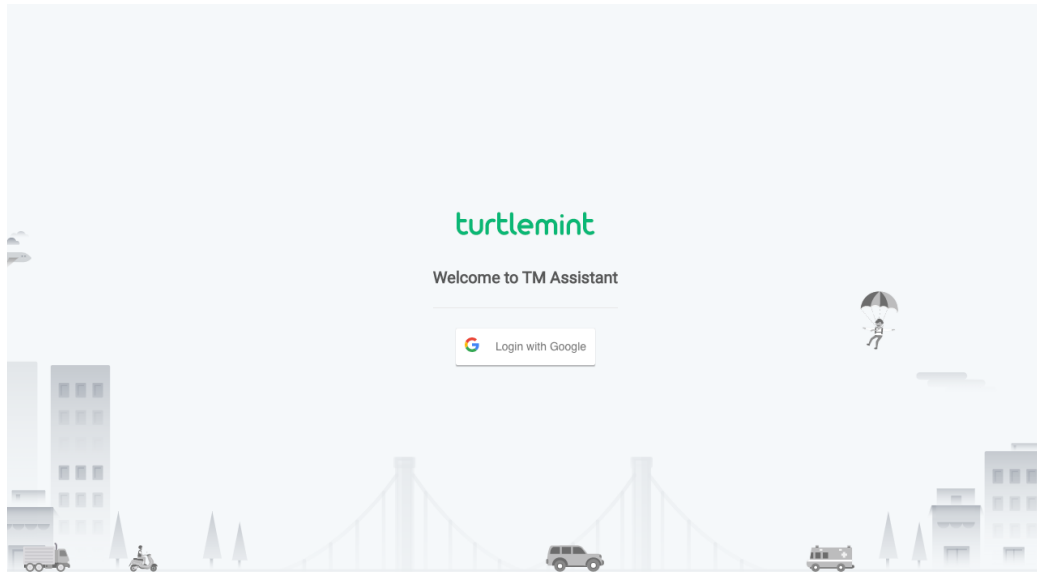


Figure 5.1: The Turtlemint assistant social auth screen

5.1.2 The features

The application provides a ton lot of features assisting the team with their updates. Here is a summary of the features we developed:

- Social authentication – Login through Google

The user can simply login via his Google account in Turtlemint assistant. The benefit being the user don't have to remember its username, password. Also, the Google only permits users with Turtlemint issued domain emails to login and proceed towards the application.

- The insurer-wise mapping data stats

There is a separate screen for the user to check which data is yet to be processed or *mapped* to the Turtlemint database. This stats are displayed insurer-wise are are extremely helpful in determining which

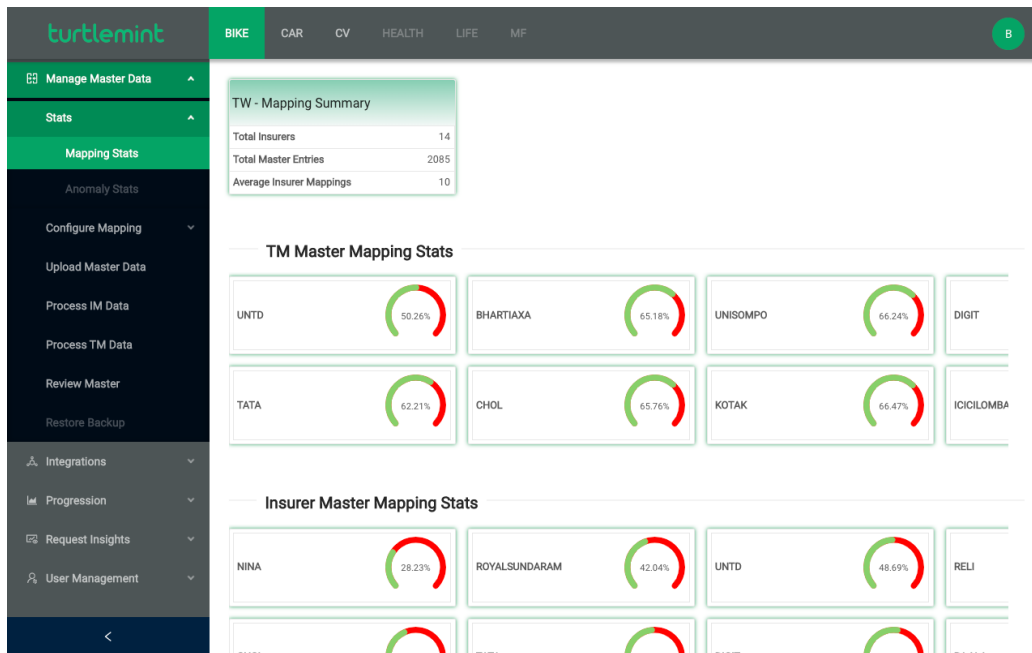


Figure 5.2: The Turtlemint assistant stats screen

insurer data needs attention and to track overall progress.

- Upload raw insurer file and calculate changes.

The user can simply upload the received file from the insurer and let the application detect all new changes received from the file. Then the user only needs to work on those new changes instead of going through all changes.

- Process all data and take suitable action

The data is distributed between various type of changes and the user can work on them individually wither by accepting or rejecting new changes. This is split across all insurers and hence multiple users can work on set of different insurers at the same time, optimizing the overall process.

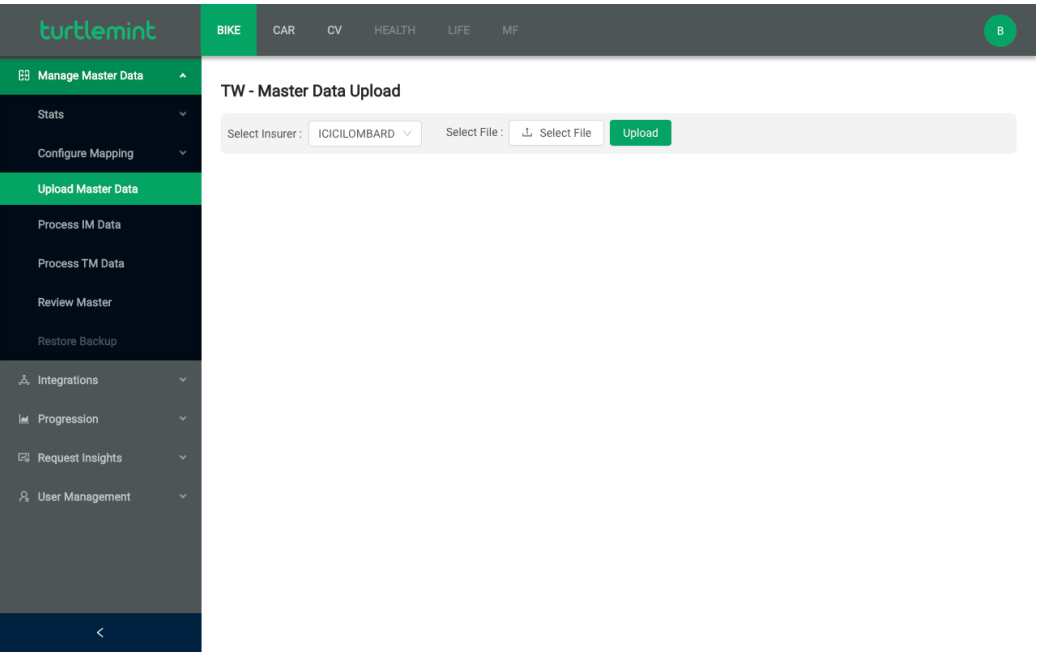


Figure 5.3: The Turtlemint assistant upload screen

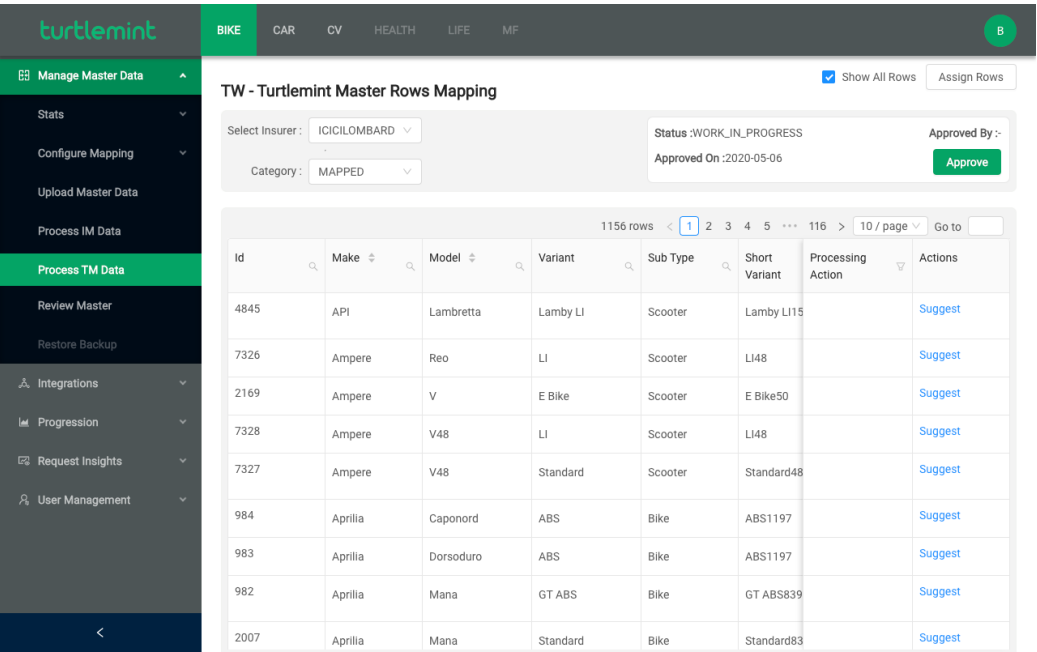


Figure 5.4: The Turtlemint assistant data screen

turtlemint BIKE CAR CV HEALTH LIFE MF B

Manage Master Data Manual Search Back

Stats ▼

Configure Mapping ▼

Upload Master Data

Process IM Data

Process TM Data

Review Master

Restore Backup

Integrations ▼

Progression ▼

Request Insights ▼

User Management ▼

TW - Current Record

Id	Cc	Fuel	Make	Model	Seating Capacity	Sub Type	Variant
7326	48	Electric	Ampere	Reo	2	Scooter	LI

Suggestions From Insurer Master

	Id	Cc	Fuel	Make	Model	Seating Capacity	Variant	Sub Type	Score
<input type="radio"/>	1315400202	50	Electric	AMPERE	REO LI	2	EBIKE	Scooter	80.12
<input type="radio"/>	1315400302	50	Electric	AMPERE	V 48 LI	2	EBIKE	Scooter	60.32
<input type="radio"/>	1315400201	50	Electric	AMPERE	REO LA	2	EBIKE	Scooter	38.69
<input type="radio"/>	1315400301	50	Electric	AMPERE	V 48 LA	2	EBIKE	Scooter	4.68
<input type="radio"/>	1315400401	50	Electric	AMPERE	ZEAL	2	EBIKE	Scooter	4.68

6 rows < 1 > 5 / page Go to

Insurer Mapping

Update Master Not Found < >

Figure 5.5: The Turtlemint assistant suggest screen

- Suggest records to be mapped with turtlemint database

The insurer data can be mapped with the Turtlemint database making them as an active entity in the insurance issuing process. This can be easily done via the application is extremely helpful since the user doesn't have to find the record's equivalent entry in Turtlemint database manually.

5.2 MintAcademy

MintAcademy is an Learning Management Software (LMS) developed as a SaaS (Software as a Service) platform for its tenants. The MintAcademy was originally written using GCP (Google Cloud Platform), GAE (Google App Engine) and NDB Storage. With the rising issues and cost reduction,

it was and still is being re-written in Django. I'm one of the core developers involved in this migration. Until now, most of the core functionality is already migrated and is expected to go live in the month of June, 2020.

The current implementation uses GCP extensively and uses all of its features. It is currently live on production server and is serving to thousands of users at the moment.