BRIGHTBUY

AN END-TO-END PROTOTYPE
ENTIRELY DESIGNED, CODED
AND DEPLOYED BY MYSELF.

HTTPS://GITHUB.COM/RAPHAEL88/IMMOPROJECT
EMAIL : RAPHAEL.NGUYENVAN@GMAIL.COM

AGENDA

MY ROLES IN THIS PORTFOLIO PROJECT

I coded the data scrapers and the Extract, Transform and Load processes from public real estate database. Everything was done in Python.

I set up an SQL Database on Azure creating the database, the tables, maintaining them while also managing few security items.

I coded the machine learning models that predict real estates price using scikit-learn library.

I coded the API using GitHub (Code versioning and repository), Render to deploy and host the API. The API is coded in python.

I designed low coded the front-end application via retool to create a prototype

I crafted the narrative behind the brand and product to align with the user-facing presentation.

EXECUTIVE SUMMARY

BrightBuy is a unique real estate analytics platform built to empower everyday buyers with the clarity and tools they need to make confident, data-driven decisions. By leveraging real transaction data, predictive modeling, and an intuitive interface, we give users an edge typically reserved for industry insiders.

The Problem

Buying property is one of life's biggest financial decisions — yet most buyers face it with uncertainty, shallow insights, and biased advice. Valuation tools are opaque, inconsistent, and seller-oriented. As a result, buyers are left wondering: "Is this a good deal?"

Our Vision

We recognize a global shift: people increasingly seek to outsmart traditional systems and step outside rigid norms. Our vision is to give everyday real estate buyers unfair clarity to outsmart the market — and the confidence to become empowered owners.

Core values

Truth, Outsmart, Empowerment

OPPORTUNITY IDENTIFICATION

- Too much open data, too little clarity
- Buyers are advised by parties in interest conflicts
- Estimates vary wildly and tools are for sellers
- Locked info behind paywall

"When I bought my flat, I had no concrete way to tell if it was a good deal." Me.

"I visited a lot, but everything seemed overpriced to me" A tester

800K

REAL ESTATE TRANSACTIONS IN 2024 IN FRANCE

50 TO 100

REAL ESTATES ADS VISITED BEFORE BUYING (FRANCE, AVG)

HYPOTHESIS

“First-time real estate buyers need clear, data-driven insights from a neutral source.”

“A significant share of them are willing to exchange their email for premium-quality analysis.“

“There is clear market space for a new kind of buyer-focused analytics service.“

PERSONA

YOUNG URBAN PROFESSIONAL COUPLE

Profile
28–35 years, working in tech/finance/marketing, no children (yet).

Motivations
Stop paying rent, invest early, prepare for family.

Constraints
Lack of knowledge, family far away, few time to invest in a research

Priorities
Proximity to public transport, vibrant neighborhood, new build or well-renovated.

Pain points
Not familiar with buying process, fear of market timing, difficulty getting financing without strong parental help.

PERSONA

SINGLE PROFESSIONAL / FIRST JOB

Profile
25–32 years, CDI just acquired, high-potential income trajectory.

Motivations
Build capital early, gain independence, pride of ownership.

Constraints
Lower borrowing capacity, lack of credit history, lack of experience, lack of knowledge

Priorities
Return on investment, best deal

Pain points
Approval from bank, competing with investors for small units, understanding hidden costs.

TECHNICAL ASSUMPTIONS (TO VALIDATE)

Key questions we're validating:

- Is the data we need accessible, consistent, and reliable?
- Can we predict prices with high accuracy?
  (Industry benchmark : 5 % - 15% error margin)
- Do we have a technically differentiating edge?

TECHNICAL VALIDATION SO FAR

Key questions we’re validating:

- ✓ Is the data we need accessible, consistent, and reliable? ** Yes **
- ✓ Can we predict prices with high accuracy?
  5–12% error margin *(and improving)*
- ✓ Do we have a technically differentiating edge?
  Future forecasting based on property condition + growth logic

UNIQUE VALUE PROPOSITION

“We give everyday real estate buyers an unfair clarity to outsmart the market…“

“… and the confidence to become sharp, empowered owners.“

MVP FEATURES

An interface that bring together :

- ✓ Real prices estimation
- ✓ Property value simulation across different quality states (old, renovated, etc.)
- ✓ Qualitative market analysis
- ✓ Price growth forecast over time (model-based)
- ✓ Transparency on model accuracy for each area
- ✓ Opportunity tracking on at least one major listing website

POST MVP

- Create a paywall system to deliver daily opportunities & AI-powered alerts
- Improve prediction performance through geo-targeted model tuning
- Work on rituals

CUSTOMER JOURNEY :

YOUNG URBAN PROFESSIONAL COUPLE

	AWARENESS	INTEREST	DESIRE	RETENTION	REFERRAL
	PROMOTION	FIRST TRY	DATA-DRIVEN VISIT	I MAKE SEVERAL VISIT	I TALK ABOUT IT
Thinking	"THIS AD IS RELATED TO MY SEARCHES"	"THERE IS SOMETHING JUST FOR PEOPLE LIKE ME. I AM JUST GIVING A TRY"	"I WILL USE THE INFORMATION DURING MY VISITS"	"EACH TIME I WILL FIND A PROPERTY ON INTERNET I WILL CHECK IT ON MY NEW TOOL !"	"HEY DO YOU KNOW THIS NEW APP ?"
Doing	User clicks on the link and discover the product.	The user tries several configurations and properties and get information he needs.	The user compare the information when he visits.	The user integrate the usage of the tool as a ritual.	If user share on social media the usage of the tool, then we improve the experience.

CUSTOMER JOURNEY :

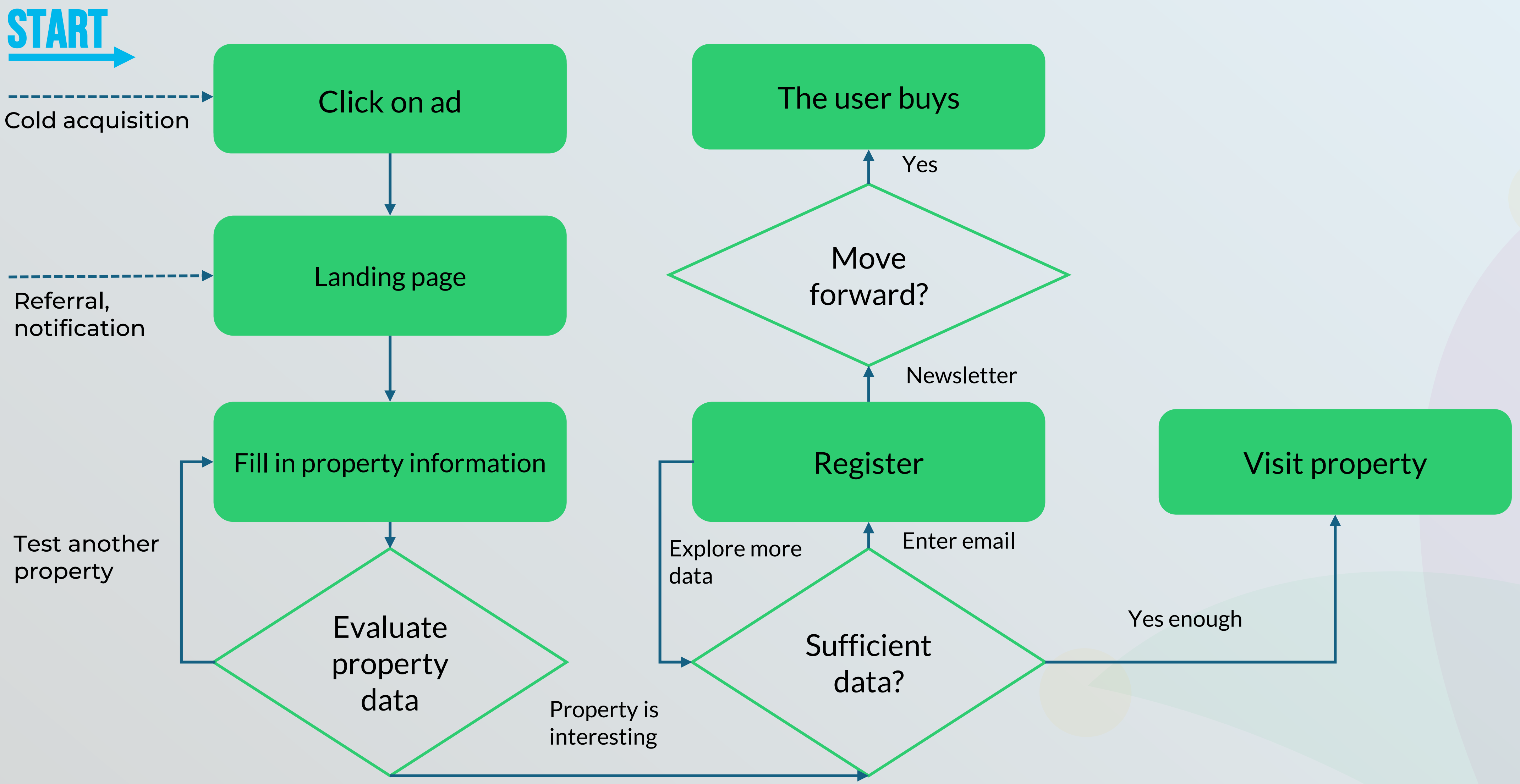
SINGLE PROFESSIONAL / FIRST JOB

	AWARENESS	INTEREST	DESIRE	ACTIVATION	ACTION
	PROMOTION	FIRST TRY	DATA-DRIVEN VISIT	I COULD BE MORE	I BUY
Thinking	“THIS AD IS RELATED TO MY SEARCHES”	“THERE IS SOMETHING JUST FOR PEOPLE LIKE ME. I AM JUST GIVING A TRY”	“I WILL USE THE INFORMATION DURING MY VISITS”	“I HAVE A UNIQUE WAY TO INVEST OTHERS WILL WANT ME TO REFER THEM.”	“I AM SCALING UP; I WANT TO HAVE FULL ANALYSIS AND LIVE OPPORTUNITIES UPDATE.”
Doing	User clicks on the link and discover the product.	The user tries several configurations and properties and get information he needs.	The user compare the information when he visits.	From the common real estate buyer, the user is turning into a small & smart investor.	The user buy the right to own the product on a combination :city + postcode.

MONETIZATION (REVENUE):

- Advertising on site (Est. 24k CHF/ Year)
- Sponsored newsletters (1’500 CHF/ Year / 1’000 Subscribers)
- Subscriptions (At least 35K potential subscriptions estimated 600K .- CHF Annually)

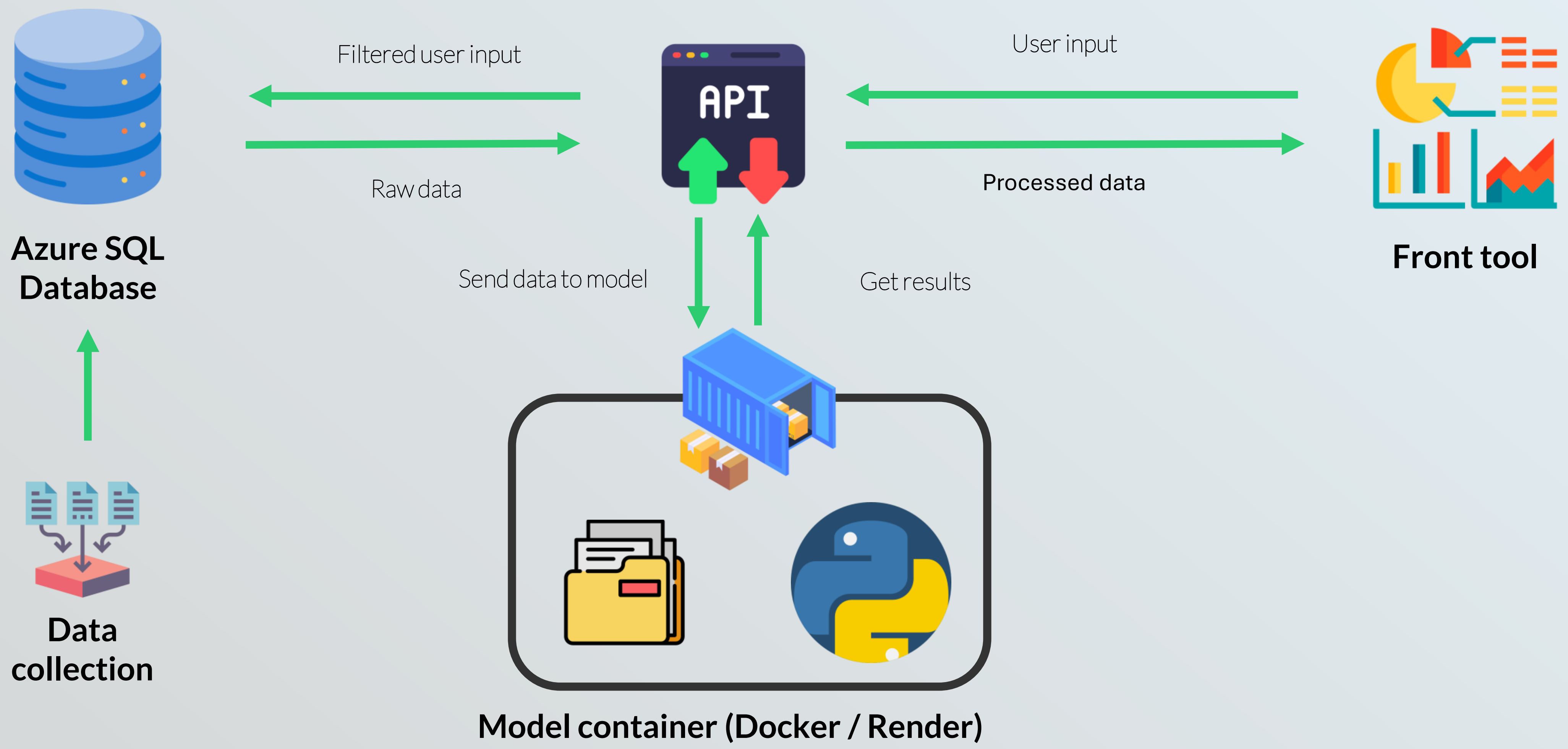
USERFLOW:


START
Cold acquisition
Click on ad
Landing page
Referral, notification
Fill in property information
Test another property
Evaluate property data
Property is interesting
Sufficient data?
Explore more data
Enter email
Register
Newsletter
Move forward?
Yes
The user buys
Yes enough
Visit property

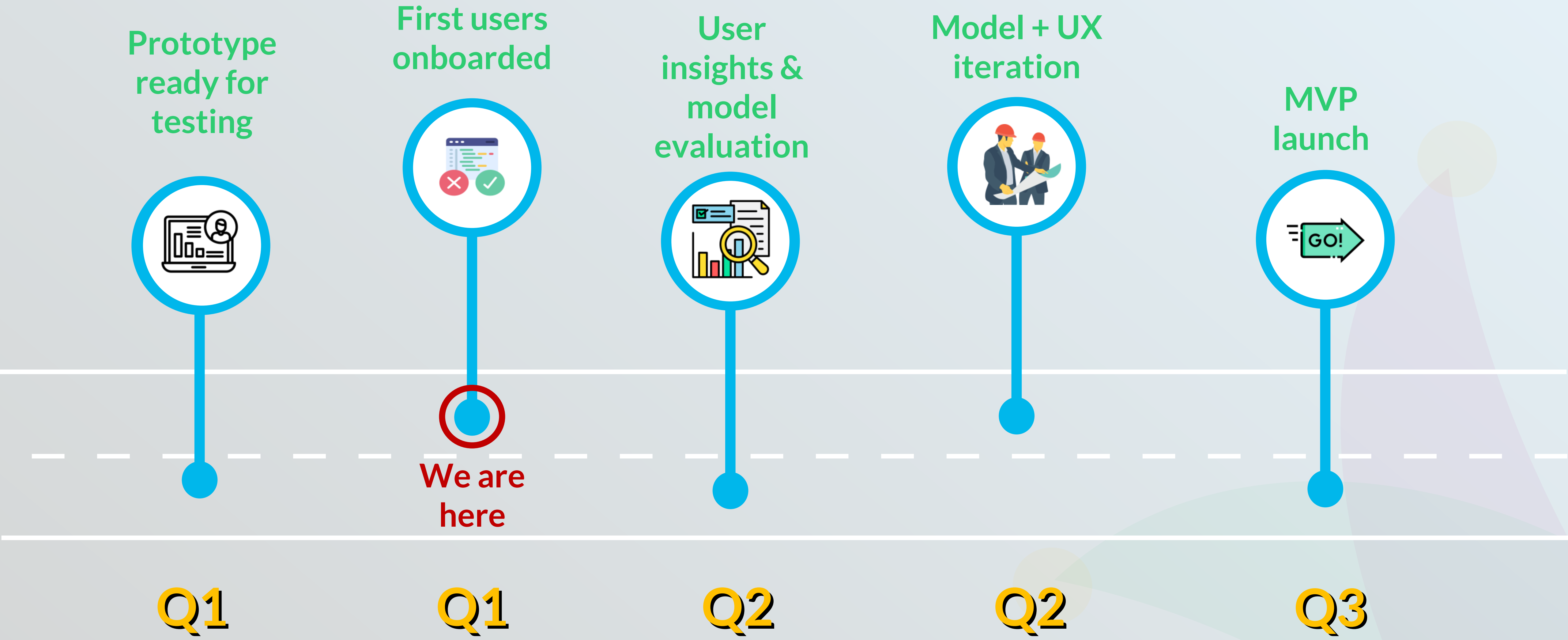
TECHNICAL STACK FOR EARLY FUNCTIONAL PROTOTYPE:

- Azure SQL Server : Stores real estate and model prediction data
- Python ML Scripts : Collect, clean, and train real estate prediction models
- Flask API : Serves predictions & data to the front-end via endpoints
- Docker & Render : Containerize and deploy the backend for testing/demo
- GitHub : Code repository for data pipeline, models, and API
- Low-code front-end (e.g. Retool / Appsmith) Build the interactive UI mockup for users and stakeholders

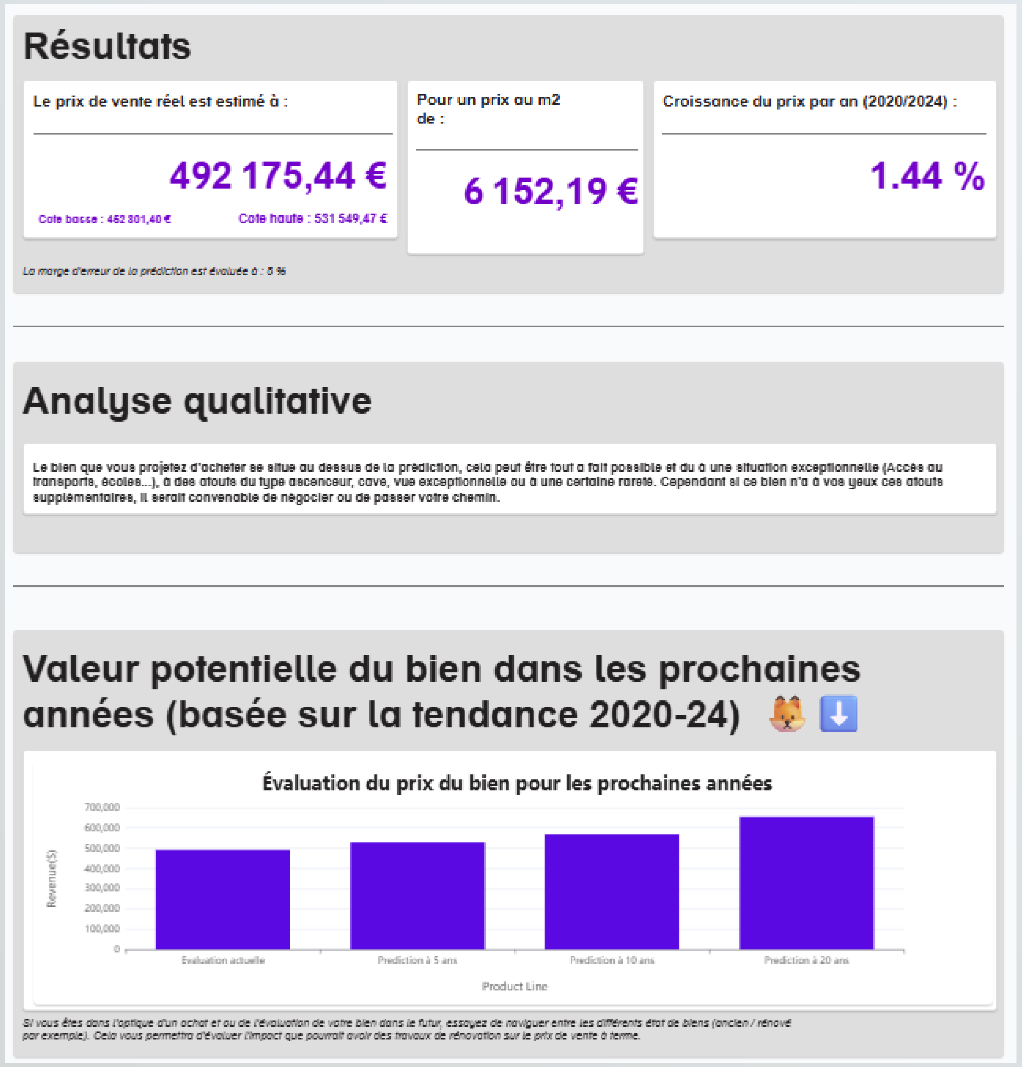
ARCHITECTURE FOR PROTOTYPE



ROADMAP:


Prototype ready for testing
Q1
First users onboarded
We are here
Q1
User insights & model evaluation
Q2
Model + UX iteration
Q2
MVP launch
Q3

PROOF OF CONCEPT (FRONT END APP VIDEO) :



Résultats

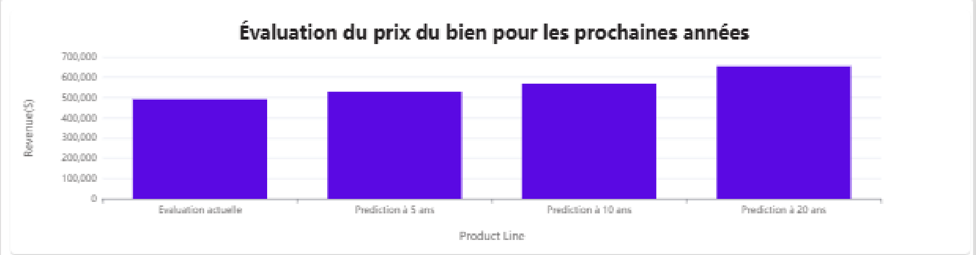
Le prix de vente réel est estimé à :	Pour un prix au m2 de :	Croissance du prix par an (2020/2024) :
492 175,44 €	6 152,19 €	1.44 %
Cote basse : 462 801,40 € Cote haute : 531 549,47 €		

La marge d'erreur de la prédiction est évaluée à : 6 %

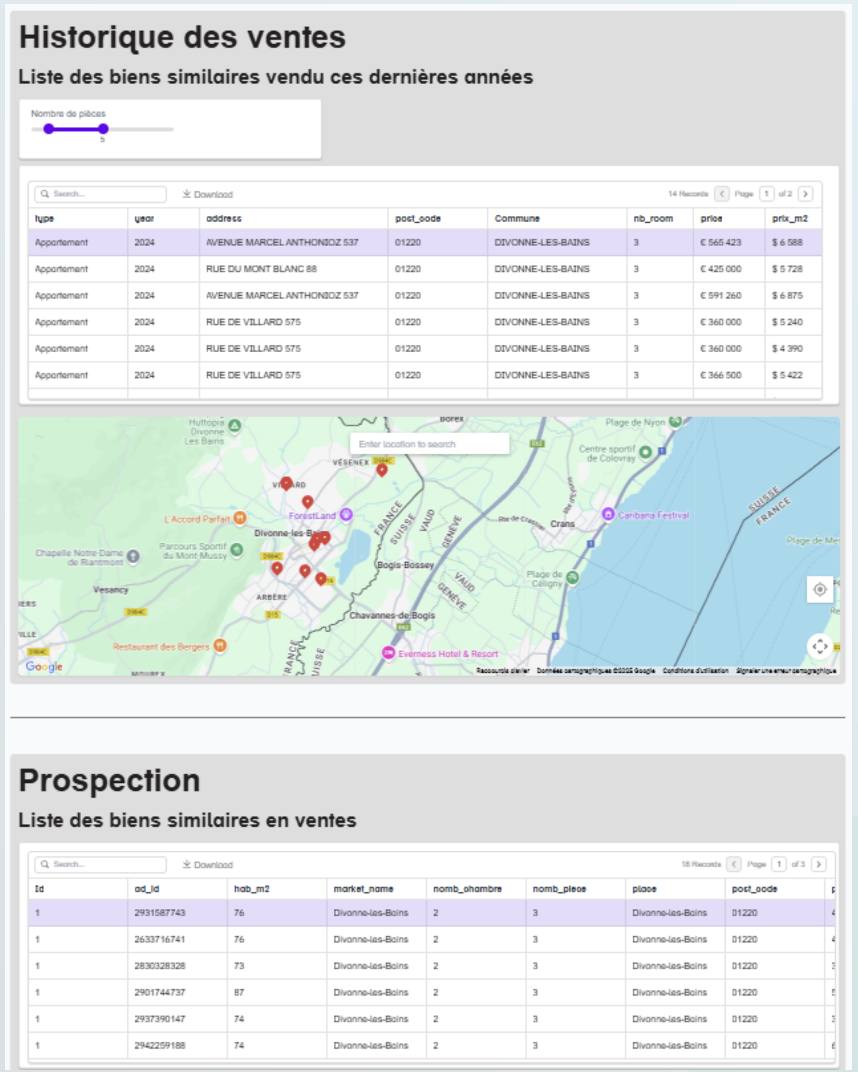
Analyse qualitative

Le bien que vous projetez d'acheter se situe au dessus de la prédiction, cela peut être tout a fait possible et du à une situation exceptionnelle (Accès au transports, écoles...), à des atouts du type ascenceur, cave, vue exceptionnelle ou à une certaine rareté. Cependant si ce bien n'a à vos yeux ces atouts supplémentaires, il serait convenable de négocier ou de passer votre chemin.

Valeur potentielle du bien dans les prochaines années (basée sur la tendance 2020-24) 🦊 ⬇️



Si vous êtes dans l'optique d'un achat et ou de l'évaluation de votre bien dans le futur, essayez de naviguer entre les différents état de biens (ancien / rénové par exemple). Cela vous permettra d'évaluer l'impact que pourrait avoir des travaux de rénovation sur le prix de vente à terme.



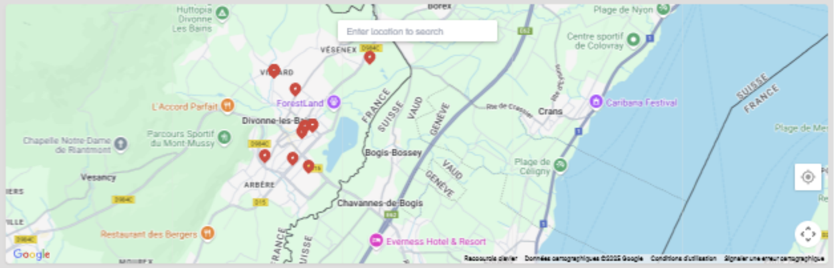
Historique des ventes

Liste des biens similaires vendu ces dernières années

Nombre de pièces 5

Search... Download 14 Records Page 1 of 2

type	year	address	post_code	Commune	nb_room	price	prix_m2
Appartement	2024	AVENUE MARCEL ANTHONIOZ 537	01220	DIVONNE-LES-BAINS	3	€ 565 423	$ 6 588
Appartement	2024	RUE DU MONT BLANC 88	01220	DIVONNE-LES-BAINS	3	€ 425 000	$ 5 728
Appartement	2024	AVENUE MARCEL ANTHONIOZ 537	01220	DIVONNE-LES-BAINS	3	€ 591 260	$ 6 875
Appartement	2024	RUE DE VILLARD 575	01220	DIVONNE-LES-BAINS	3	€ 360 000	$ 5 240
Appartement	2024	RUE DE VILLARD 575	01220	DIVONNE-LES-BAINS	3	€ 360 000	$ 4 390
Appartement	2024	RUE DE VILLARD 575	01220	DIVONNE-LES-BAINS	3	€ 366 500	$ 5 422



Prospection

Liste des biens similaires en ventes

Search... Download 18 Records Page 1 of 3

Id	ad_id	hab_m2	market_name	nomb_chambre	nomb_piece	place	post_code
1	2931587743	76	Divonne-les-Bains	2	3	Divonne-les-Bains	01220
1	2633716741	76	Divonne-les-Bains	2	3	Divonne-les-Bains	01220
1	2830328328	73	Divonne-les-Bains	2	3	Divonne-les-Bains	01220
1	2901744737	87	Divonne-les-Bains	2	3	Divonne-les-Bains	01220
1	2937390147	74	Divonne-les-Bains	2	3	Divonne-les-Bains	01220
1	2942259188	74	Divonne-les-Bains	2	3	Divonne-les-Bains	01220

SUCCESS METRICS FOR TEST:

Metrics 1 : Users finds relevant information 0 - 10

Metrics 2 : Users can produce arguments for negotiation 0 - 10

Metrics 3 : Users feels empowered and confident 0 - 10

Metrics 4 : Users want more information and are ready to give their email address 0 – 10

Metrics 5 : Users recognize a here a tool for investing 0 - 10

SUCCESS METRICS FOR V1 :

Metrics 1 : Number of email registered (10000), newsletter subscriptions (1000), Subscription sold (500)

Metrics 2 : Total number of visits, Bounce rate, Traffic sources, Time on page, Total numbers of prediction performed

BUDGET FOR TEST:

Workstream / Role	Manpower (D)	Daily Rate (CHF)	Fix costs	Total
Azure SQL Server / **Engineering**	1	800.-	Free tier	800.-
Data collection and Machine learning / **Engineering**	2	800.-	No	1600.-
API (Flask) / **Engineering**	2	800.-	No	1600.-
Docker / Render / **Engineering**	1	800.-	Free tier	800.-
Github / **Engineering**	0.1	800.-	Free tier	100.-
Low code interface designer / **UX/UI Design**	5	800.-	Free tier	4000.-
Testing and customer return / **Product manager**	1	800.-	No	800.-
Product Management	5	800.-	No	4000.-
TOTAL				16900.-

BUDGET FOR V1.00 (DEPEND ON TESTING RESULTS) :

Workstream / Role	Manpower (D)	Daily Rate (CHF)	Fix costs	Total
Azure SQL Server / **Engineering**	TBD	800.-	1’200.-	TBD
Data collection and Machine learning / **Engineering**	TBD	800.-	50.- per post code & city	TBD
API (Flask) / **Engineering**	TBD	800.-	No	TBD
Docker / Render / **Engineering**	TBD	800.-	600.-	TBD
Github / **Engineering**	TBD	800.-	Free tier	TBD
Low code interface designer / **UX/UI Design**	TBD	800.-	Free tier	TBD
CRM / Email Automation / **Customer team**	TBD	800.-	No	TBD
Analytics & Tracking Setup / **Product analyst team**	TBD	800.-	No	TBD
Testing and customer return / **Product manager**	TBD	800.-	No	TBD
Product Management / Marketing	TBD	800.-	No	TBD
TOTAL				TBD

GO TO MARKET:

Launch owner	RNV
Target persona	Young Urban Professional Couple / Single Professional / First Job
Monetization Strategy	Ads, newsletters and subscriptions
Marketing strategy	Google Ads, word of mouth
Campaign Effectiveness	Q3 – Google analystics
Customer Adoption	Customer Support - FAQs, Communication through Video and Static Messages on App & through Emails Tutorials
Support & Maintenance	Engineering team

THANK YOU SO MUCH FOR YOUR TIME

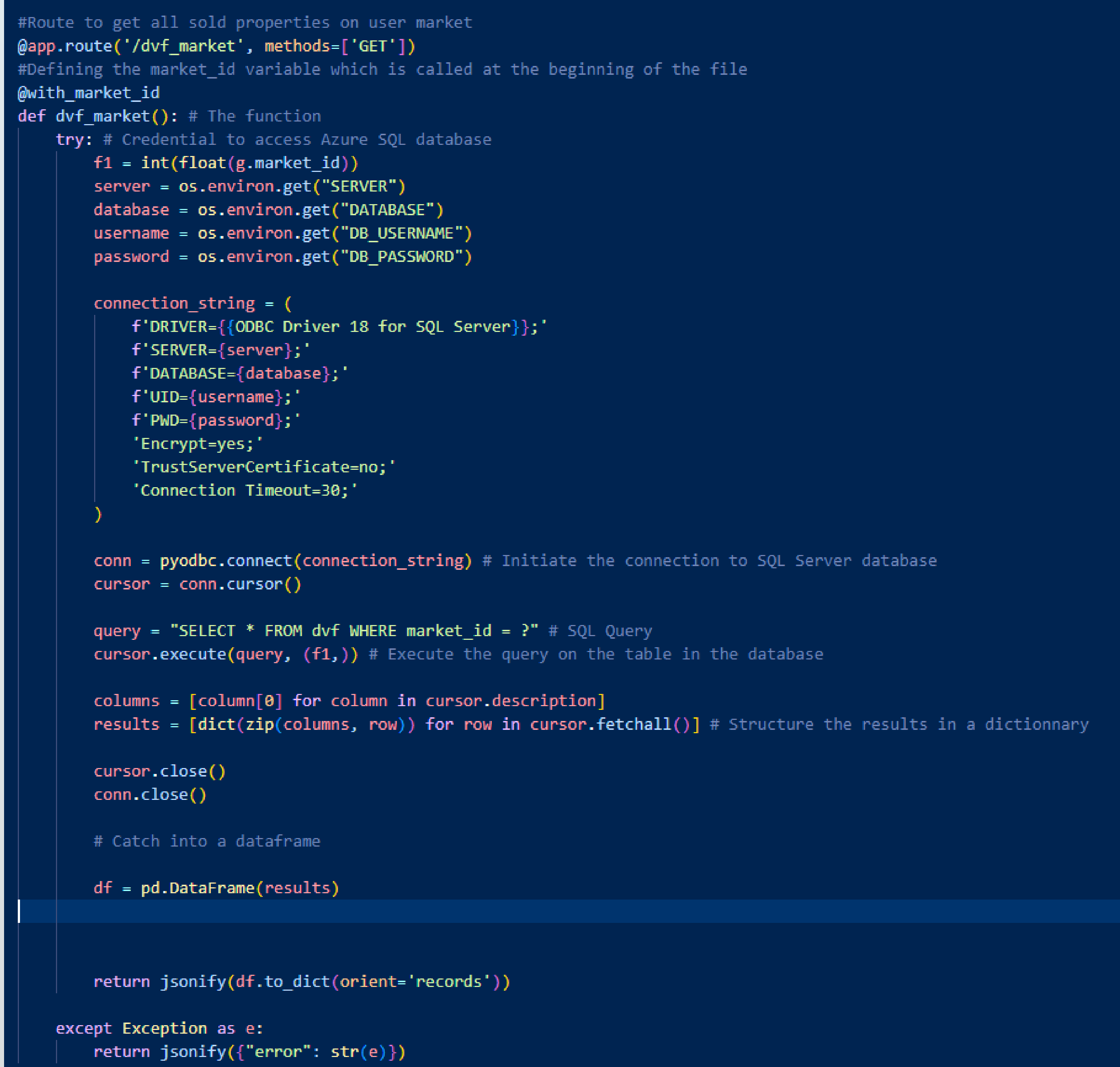
CODE SAMPLES: API

AN API ROUTE TO RETRIEVE DATA FROM THE SQL DATABASE

This is a code sample that I write to bring data from my SQL database to my front-end application.

I choose to fetch the results as a dictionary because my visualization is a table.

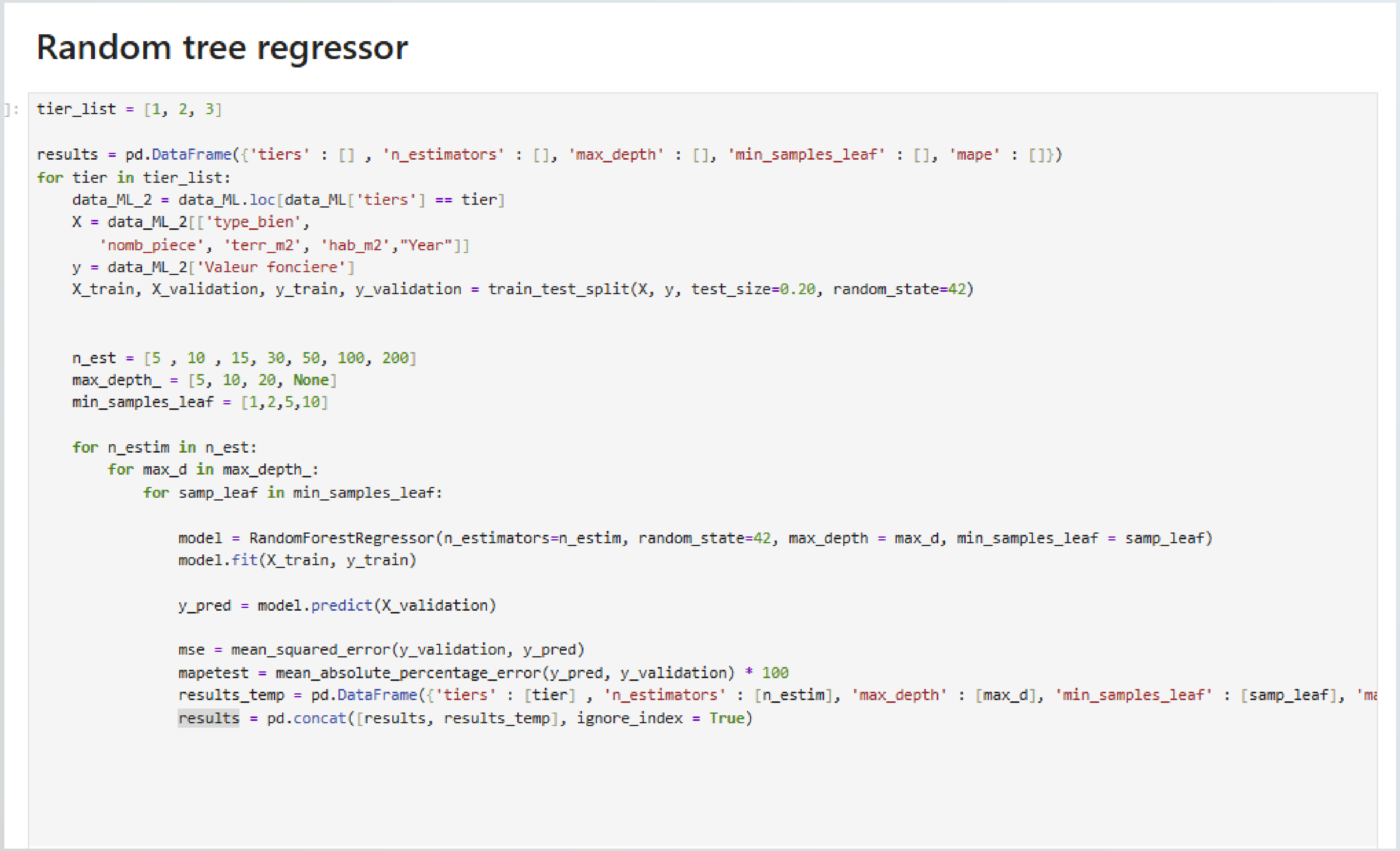
For security purposes all credential are set in render environment and are not hard coded.



```
#Route to get all sold properties on user market
@app.route('/dvf_market', methods=['GET'])
#Defining the market_id variable which is called at the beginning of the file
@with_market_id
def dvf_market(): # The function
    try: # Credential to access Azure SQL database
        f1 = int(float(g.market_id))
        server = os.environ.get("SERVER")
        database = os.environ.get("DATABASE")
        username = os.environ.get("DB_USERNAME")
        password = os.environ.get("DB_PASSWORD")

        connection_string = (
            f'DRIVER={{ODBC Driver 18 for SQL Server}};'
            f'SERVER={server};'
            f'DATABASE={database};'
            f'UID={username};'
            f'PWD={password};'
            'Encrypt=yes;'
            'TrustServerCertificate=no;'
            'Connection Timeout=30;'
        )

        conn = pyodbc.connect(connection_string) # Initiate the connection to SQL Server database
        cursor = conn.cursor()

        query = "SELECT * FROM dvf WHERE market_id = ?" # SQL Query
        cursor.execute(query, (f1,)) # Execute the query on the table in the database

        columns = [column[0] for column in cursor.description]
        results = [dict(zip(columns, row)) for row in cursor.fetchall()] # Structure the results in a dictionnary

        cursor.close()
        conn.close()

        # Catch into a dataframe

        df = pd.DataFrame(results)



        return jsonify(df.to_dict(orient='records'))

    except Exception as e:
        return jsonify({"error": str(e)})
```

CODE SAMPLES: MACHINE LEARNING

HOW TO PREDICT THE HOUSE AND APARTMENT PRICES



Random tree regressor

```
tier_list = [1, 2, 3]

results = pd.DataFrame({'tiers' : [] , 'n_estimators' : [], 'max_depth' : [], 'min_samples_leaf' : [], 'mape' : []})
for tier in tier_list:
    data_ML_2 = data_ML.loc[data_ML['tiers'] == tier]
    X = data_ML_2[['type_bien',
       'nomb_piece', 'terr_m2', 'hab_m2',"Year"]]
    y = data_ML_2['Valeur fonciere']
    X_train, X_validation, y_train, y_validation = train_test_split(X, y, test_size=0.20, random_state=42)


    n_est = [5 , 10 , 15, 30, 50, 100, 200]
    max_depth_ = [5, 10, 20, None]
    min_samples_leaf = [1,2,5,10]

    for n_estim in n_est:
        for max_d in max_depth_:
            for samp_leaf in min_samples_leaf:

                model = RandomForestRegressor(n_estimators=n_estim, random_state=42, max_depth = max_d, min_samples_leaf = samp_leaf)
                model.fit(X_train, y_train)

                y_pred = model.predict(X_validation)

                mse = mean_squared_error(y_validation, y_pred)
                mapetest = mean_absolute_percentage_error(y_pred, y_validation) * 100
                results_temp = pd.DataFrame({'tiers' : [tier] , 'n_estimators' : [n_estim], 'max_depth' : [max_d], 'min_samples_leaf' : [samp_leaf], 'ma
                results = pd.concat([results, results_temp], ignore_index = True)
```

This is a code sample of a random tree regressor exploration.

I coded a nested fine-tuning to evaluate different fine parameters all in one place.

For simplicity purposes the model used for prototype is a linear regression but random tree regressor is interesting and promising.