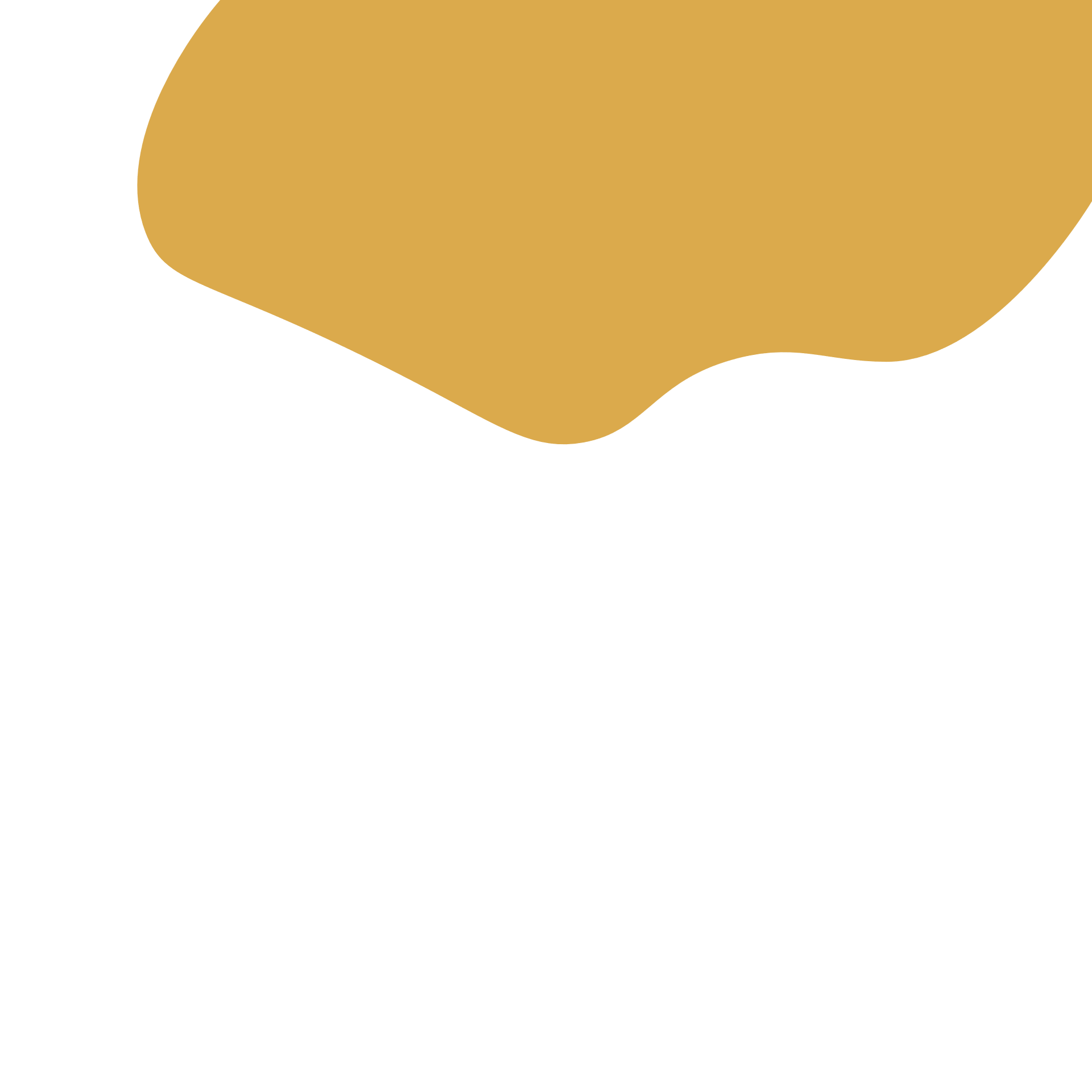
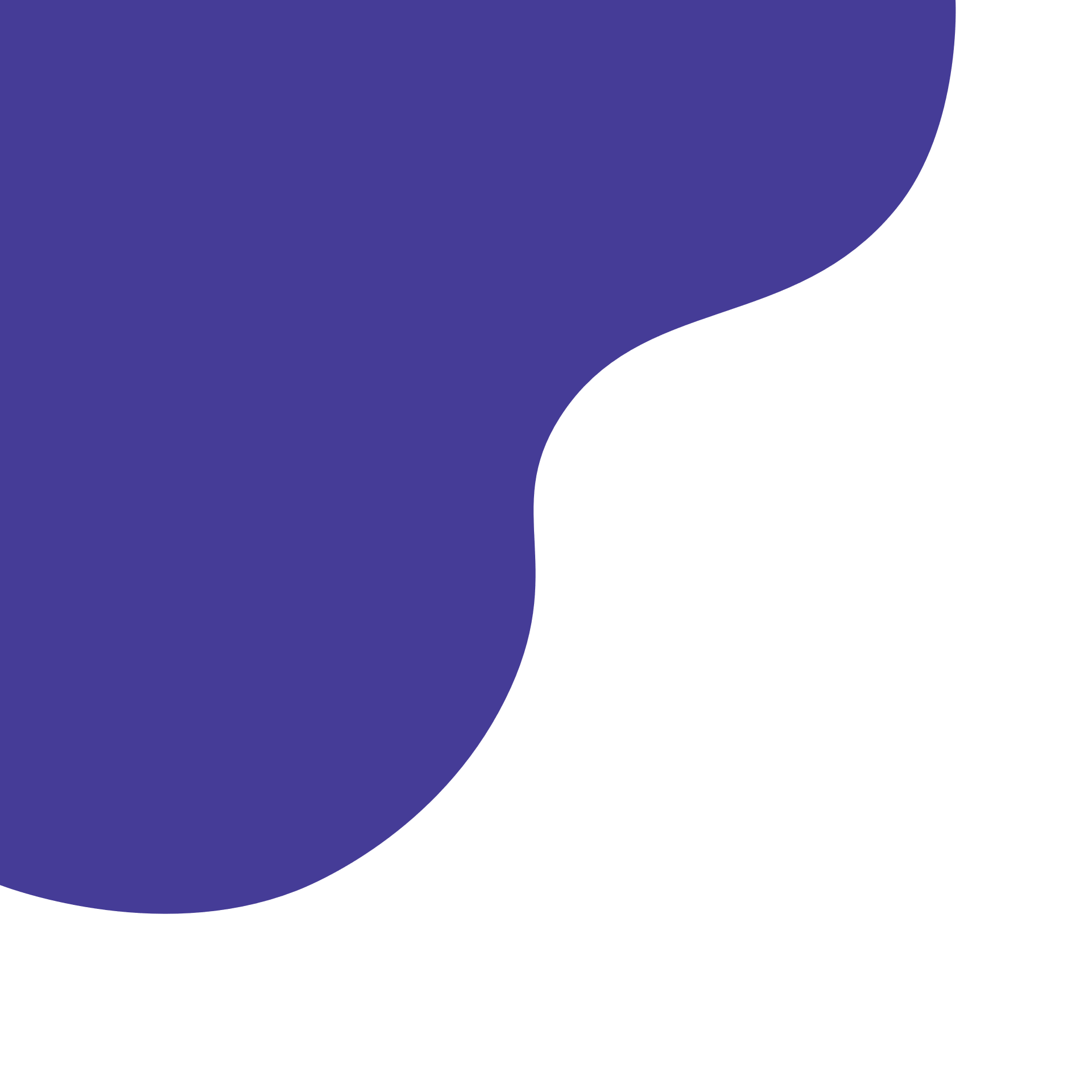
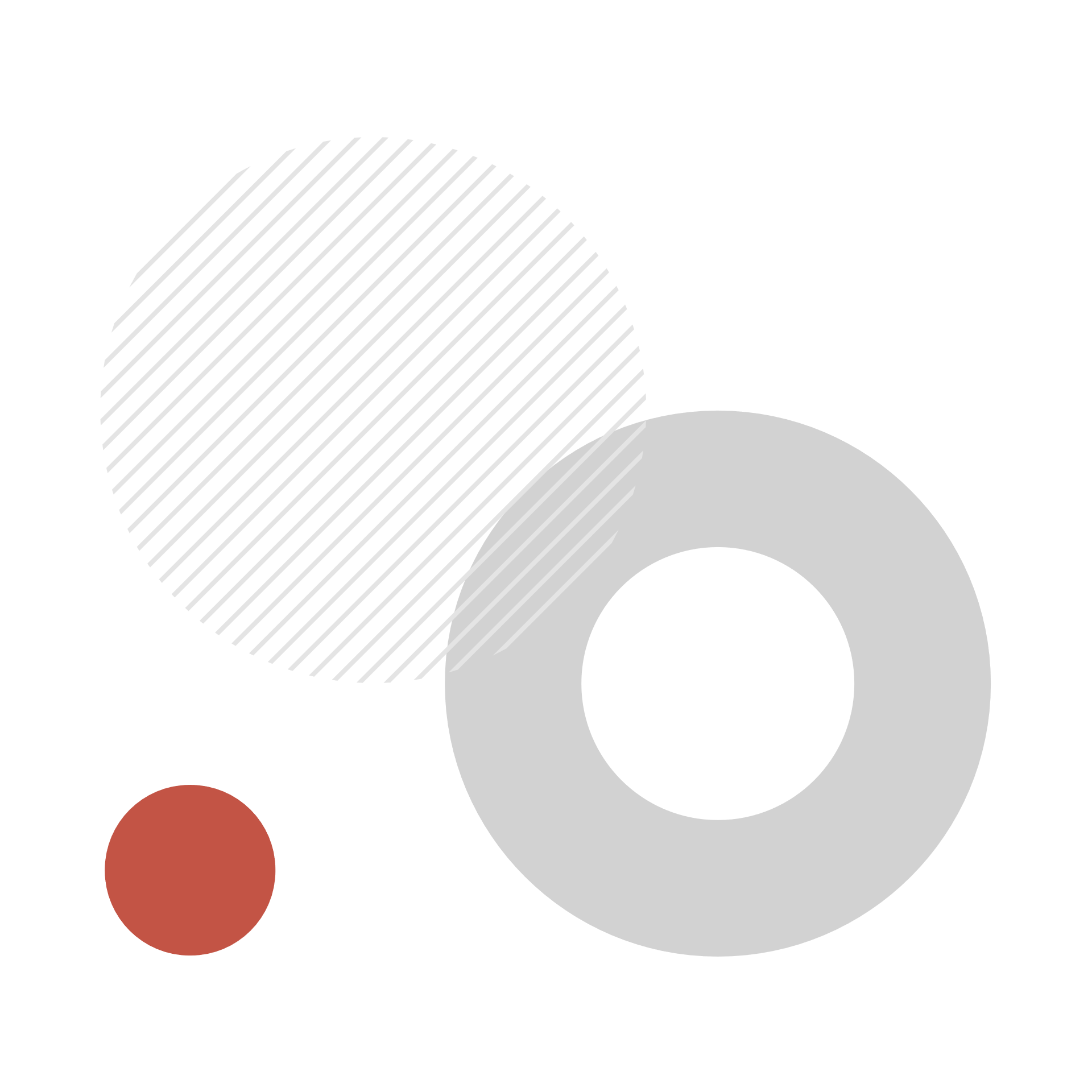
**Heart Failure**

**Case Study**

*Documentation*

****Prepared by:

**Muhammad Gaber**

**Abdalla Hamid**

**Abdelrahman Zakaa**

**Muhammad Soliman**

**Mahmoud Saad**

**Heart Failure Project Analysis**

*Documentation*

* **Project Proposal**

Heart disease is a major health concern and a leading cause of death worldwide, often going unnoticed until it reaches a critical stage. This project aims to analyze and predict heart disease at an early stage, helping individuals recognize potential symptoms and take preventive measures to improve their health and longevity. Using a real-life dataset from [Kaggle](https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction?resource=download), we will explore key risk factors and patterns associated with heart disease. Python and Excel will be used for data processing, analysis, and prediction, while Tableau will help visualize insights through interactive dashboards. By identifying important trends and risk factors, this project seeks to contribute to early detection and prevention efforts, ultimately promoting better heart health.

* **Project Plan**

|  |  |
| --- | --- |
| **Month** | **Working On** |
| January | Excel Analysis |
| February | Python Analysis & Predictions |
| March | Tableau visualization |
| April | Polishing & setting the last touches |

* **Task Assignment & Roles**

Excel **⟶** Abdelrahman Zakaa & Abdalla Hamed

Python **⟶** Muhammad Gaber & Muhammad Soliman

Tableau **⟶** Muhammad Gaber & Mahmoud Saad

* **Risk Assessment & Mitigation Plan**
* Balancing Time Between Work and Tasks **⟶** Saving 1 Hour per day for finishing our tasks
* Working with Ramadan hours **⟶** Starting on our tasks after i’sha
* Finding time that suits everyone **⟶** creating zoom meetings to save travel time
* **KPIs**

**1. Data Processing & Cleaning (January - Excel Analysis)**

*Goal:* Ensure high-quality, structured data for analysis

**KPIs:**

* **Data Completeness:**
  + *Target:* 100% of dataset loaded and structured correctly
  + *Metric:* Percentage of missing values handled (should be <5%)
* **Data Accuracy & Integrity:**
  + *Target:* Ensure that cleaned data matches the original dataset without errors
  + *Metric:* Number of inconsistencies or duplicates removed
* **Exploratory Insights Generated:**
  + *Target:* Identify at least 3 key trends/patterns using pivot tables
  + *Metric:* Number of meaningful insights derived from Excel analysis

**2. Data Analysis & Machine Learning (February - Python Analysis & Predictions)**

*Goal:* Build predictive models and uncover insights

**KPIs:**

* **Correlation & Feature Importance:**
  + *Target:* Identify **the top 3-5 factors** influencing heart failure
  + *Metric:* Correlation matrix and statistical significance tests (p-value <0.05)
* **Anomalies & Outlier Detection:**
  + *Target:* Identify and explain anomalies in the dataset
  + *Metric:* Number of outliers detected and handled (e.g., using IQR or Z-score methods)
* **Comparison of Different Patient Groups:**
  + *Target:* Compare heart failure trends across different demographics (age, gender, etc.)
  + *Metric:* Number of insights on variations across patient subgroups
* **Feature Engineering Efficiency:**
  + *Target:* Identify at least 3 key features that improve model accuracy
  + *Metric:* Feature importance ranking and contribution to model performance
* **Descriptive Statistics & Trends:** 
  + *Target:* Calculate key statistics (mean, median, standard deviation) for each feature
  + *Metric:* Number of variables analyzed and their impact on health outcomes
* **Exploratory Data Analysis (EDA):**
  + *Target:* Create 5+ visualizations (scatter plots, heatmaps, etc.)
  + *Metric:* Number of insights derived from Python visualizations

**3. Visualization & Dashboard Development (March - Tableau Visualization)**

*Goal:* Create meaningful visualizations for decision-making

**KPIs:**

* **Dashboard Completion:**
  + *Target:* Build 3+ interactive dashboards for analysis
  + *Metric:* Number of visualizations created in Tableau
* **Data Storytelling Effectiveness:**
  + *Target:* Insights should be clear and actionable
  + *Metric:* Number of insights that drive decisions or change in understanding

**4. Project Completion & Refinement (April - Final Touches & Reporting)**

*Goal:* Finalize project, document findings, and ensure usability

**KPIs:**

* **Final Report Completion:**
  + *Target:* Submit a comprehensive project report covering all phases
  + *Metric:* Number of pages covering analysis, insights, and recommendations
* **Model & Dashboard Validation:**
  + *Target:* Ensure 100% functionality with no critical errors
  + *Metric:* Number of bug fixes and improvements made before deployment
* **Project Delivery on Time:**
  + *Target:* Complete all phases by end of April
  + *Metric:* Percentage of milestones completed on schedule

**Requirement Gathering**

* **Stakeholder**

This is an independent project with no foreign partner, the dataset used is published publicly for anyone to use, The final result of our work will be published on our portfolio.

* **User Stories & Cases**

Our dataset has 12 Features with one of them being the target. We will be asking a question linking each of the other 11 feature to the target to see how do they affect it and how much of an effect it is.

**\*Examples**

1. Which gender suffers from heart disease more?
2. What heart rate range can we be concerned by the probability of having a heart attack?
3. What range of cholesterol should I start to worry about for my health?

Answering these questions using analytical means and visualization will help us get insights to understand the target more and predict it beforehand.

* **Functional Requirements**

We Will be using excel to create pivot tables, ready the columns before exporting to python and tableau and using built-in functions to get insights.

We will use and focus on Python as our main tool that will prepare and clean the data, answer our questions, make graphs and charts and predict the target behavior for the future.

Finally, we will use tableau to create visualizations and dashboards to easier understand our data and our findings

* **Non-Functional Requirements**

Maintaining our performance through the whole project

Finding compatibility & availability between all team members

**System Analysis & Design**

* **Problem Statement & Objectives**

As we stated before, the objective of this project is to analyze and predict heart disease at an early stage, helping individuals recognize potential symptoms and take preventive measures to improve their health and longevity.

* **Data Flow**

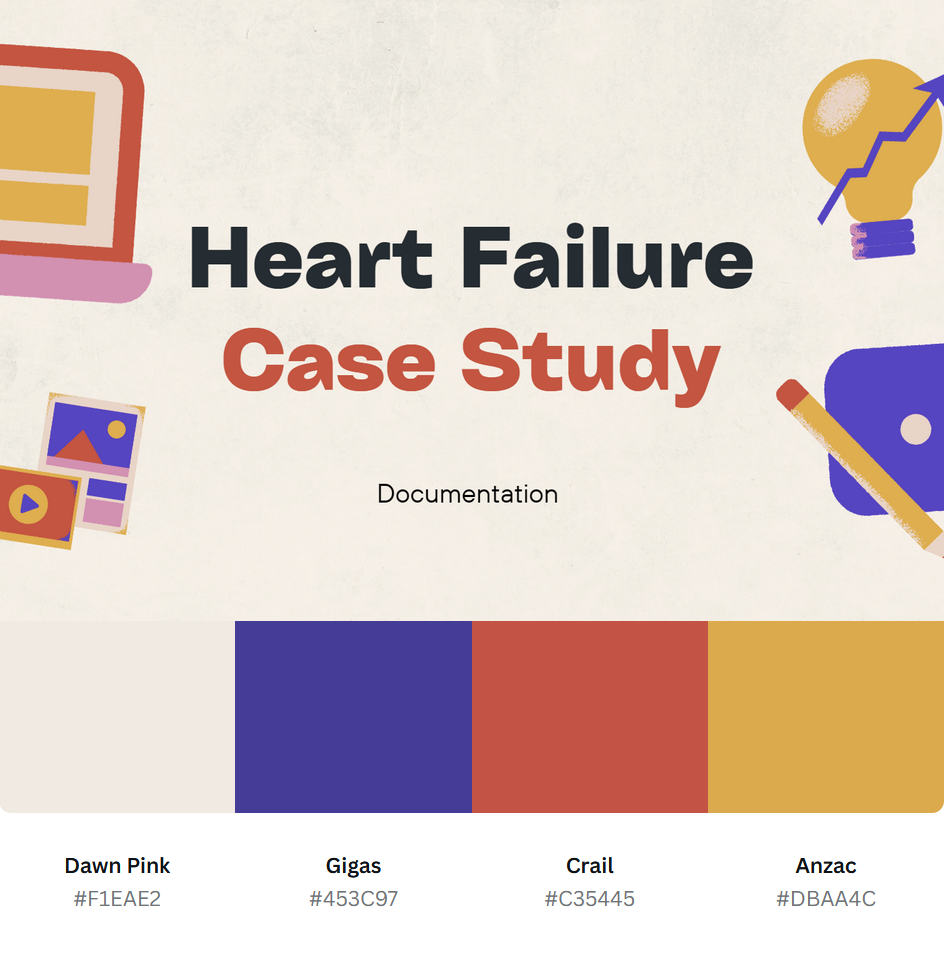
**1. Data Source:**

* Download Dataset from Kaggle (Heart Failure Prediction)

**2. Data Processing Steps:**

* **Excel:**
  + Load dataset
  + Create pivot tables for basic analysis
  + Basic Aggregations
* **Python:**
  + Data Cleaning & Preprocessing
  + Handle Missing Values, Outliers
  + Feature Engineering
  + Exploratory Data Analysis (EDA)
  + Graphs & Charts
  + Identify Trends, Correlations
  + Generate Insights and Analytical Questions
  + Build Machine Learning Models
  + Train & Test Models
  + Evaluate Performance
* **Tableau:**
  + Load Cleaned Data
  + Create Interactive Dashboards
  + Generate Visualizations (Bar, Line, Heatmaps, etc.)
  + Present Insights Visually

**3. Publishing Our Results**

* **UI/UX Design & Prototyping**