# Data Analysis Project

## Data Analyst: Arhan Khan

## Client/Sponsor: Google Data analytics Capstone

## Purpose:

*The purpose of this project is to explore how environmental factors like temperature, humidity, total dissolved solids (TDS), and pH levels impact the growth duration of lettuce. The goal is to derive data-driven insights that can inform better agricultural practices and optimize growth conditions for future harvests.*

## Scope / Major Project Activities:

*What are the major parts of this project? List out the high-level steps, activities, or stages of the project, and give a brief description for each.*

| Activity | Description |
| --- | --- |
| Dataset Selection and  Import | Selected the *Lettuce Growth Days* dataset from Kaggle and loaded it into Sheets |
| Data cleaning | Handled missing values and non-numeric data using formulas (e.g., ARRAYFORMULA, VALUE) |
| Calculations,analysis,visualization and result(Interpretation) | Created key charts, calculated various values and shared data driven insights |

## This project does not include:

*Specify the things that this project isn’t responsible for doing (out of scope). For instance, “this project does not involve a summation of 2019 data analysis”*

* Analysis outside the given dataset scope
* Crop yield or economic impact forecasting
* Weather forecasting or real-time sensor data

## Deliverables:

*A specific list of things that your project will deliver.*

| Deliverable | Description/ Details |
| --- | --- |
| Cleaned Dataset | Lettuce growth dataset with numeric conversions and usable formats |
| Summary and Analysis sheet with visualization | Table with key statistical metrics and Pearson correlations between all relevant pairs of variables and 4 key charts. |
| Project Documentation | Complete SOW (including purpose, scope, deliverables, milestones, insights, etc.) |

## Schedule Overview / Major Milestones:

*The expected schedule for the project. This can be defined by milestones (e.g. “all data is cleaned and processed”), periods of time (“Week 1 / Week 2”), or other ways based on the needs of the project.*

| Milestone | Expected Completion Date | Description/Details |
| --- | --- | --- |
| *Dataset imported and cleaned* | *24/06/2025* | *Data loaded and formatted in Sheets* |
| *Summary Stats & Correlation Done* | *26/06/2025* | *Computed statistical summaries and correlation matrix and created and reviewed all graph in the visualization tab* |
| *Final SOW documented* | *26/06/2025* | *This document drafted and reviewed* |

## \*Estimated date for completion:

*It will be completed by 26th june 2025*