Improving Cost Estimation and Project Planning for Contract-Based Furniture Projects

A Mid-Term report for the BDM capstone Project

Submitted by

Name: THAKUR VISHALKUMAR VINOD

Roll number: 23f2004320

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

Declaration Statement

I am working on a Project titled "**Improving Cost Estimation and Project Planning for Contract-Based Furniture Projects**". I extend my appreciation to **The Furniture Master**, for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered from primary sources and carefully analyzed to assure its reliability.

Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through thorough analytical procedures.

I am dedicated to adhering to the principles of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I understand that all recommendations made in this project report are within the context of the academic project taken up towards course fulfillment in the BS Degree Program offered by IIT Madras. The institution does not endorse any of the claims or comments.

Signature of Candidate: (**Digital Signature**)

Name: THAKUR VISHALKUMAR VINOD

Date: sep 08, 2025

1 Executive Summary

Project Title: Improving Cost Estimation and Project Planning for Contract-Based Furniture Projects

The Furniture Master Interiors is a custom furniture company in Pune that works on contract projects for homes, offices, and institutions. Every project is unique because it is designed to fit the client's needs, space, and style. Over the years, the company has built a good name for quality work and creative designs. However, as the number of projects has grown, it has faced new challenges in managing costs and timelines.

The first big problem is that the cost estimated at the quotation stage is often lower than the actual cost. This means the company earns less profit than expected. The second problem is that many projects are delayed. Delays happen due to late delivery of materials, shortage of skilled workers, or changes in the design after work has started. These problems reduce profits and also affect customer satisfaction.

To address these issues, this project will use a data-driven approach. Primary data was collected manually over a 6-month period (January 2025 – June 2025) from past projects. The data includes quotations, bills, purchase records, site reports, and team feedback. By studying this information, we can find patterns that cause cost overruns and delays.

Based on these findings, we will create simple tools like a cost estimation sheet, a material planning chart, and a labor scheduling tracker. These tools will help the company give accurate quotations, plan resources better, and complete projects on time. This will lead to happier customers, better profits, and a stronger position in the markets.

2 Proof of Originality of Data

The data used in this project was collected directly from The Furniture Master Interiors, ensuring its authenticity and relevance. Over six months (January–June 2025), primary records such as project quotations, final bills, and purchase orders were compiled manually. Each entry was cross-checked against physical invoices, material delivery slips, and labor attendance logs to confirm accuracy.

To maintain transparency, the dataset includes only 25 completed projects, excluding ongoing or disputed contracts. Key metrics quoted vs. actual costs, material types (e.g., plywood, MDF) and delay reasons were extracted from these records.

To validate the originality of this primary data, the following supporting materials have been compiled:

- Signed letter from the business owner confirming data access and usage rights.
- Site images showing project stages.
- Video recording (in English) of a discussion with the owner, explaining typical project challenges.
- Scanned copies of manual records (e.g., handwritten purchase logs, labor register, material purchase lists, etc).

All of these resources have been compiled and securely shared via the following Google Drive link: https://drive.google.com/drive/folders/1Qs9xWRGmQ757cVd6joQGi8l_xCGL-VxJ

3 Metadata and Descriptive Statistics

3.1 Metadata

This dataset contains information on 25 contract-based furniture projects completed by The Furniture Master Interiors. Each project is recorded with 11 key details, including project ID, type of project, size in square feet, materials used, quoted and actual costs, estimated and actual timelines, cost overrun percentage, delay in days, and the main issues faced.

The data is a mix of numbers and text. For example, numerical fields such as project size, quoted cost, actual cost, and days help measure performance, while text fields like project type, material choice, and key issues give context to why costs or delays occurred. On average, projects had cost overruns and delays, which highlight challenges in planning and execution.

The data was collected manually over six months (January–June 2025) from quotations, purchase records, site reports, and team feedback. It is stored in Excel (.xlsx) format for easy analysis and visualization.

Variable	Description	Data Type
Project ID	Unique identifier for each project (e.g., P001, P002)	Categorical

Project Type	Furniture category (e.g., Modular Kitchen, Wardrobe Set)	Categorical
Size (sq. ft.)	Total area of the project	Numerical
Material Choice	Primary materials used (e.g., Plywood + Laminate, MDF + Veneer)	Categorical
Quoted Cost (₹)	Initial estimated cost provided to the client	Numerical
Actual Cost (₹)	Final incurred cost	Numerical
Estimated Days	Planned project duration	Numerical
Actual Days	Real completion time	Numerical
Cost Overrun (%)	Percentage by which actual cost exceeded the quote	Numerical
Delay (days)	Additional days taken beyond the estimated timeline	Numerical
Key Issues	Primary reasons for delays/cost overruns (e.g., supplier delays, labor shortages)	Text

Link to Data -

https://docs.google.com/spreadsheets/d/1DjugC00dXkwND3rMDY_iP1sH1PCafZdJZ2X9P9N_I08/edit?usp=sharing .

3.2 Descriptive Statistics

Metric	Size (sq. ft.)	Quoted Cost (₹)	Actual Cost (₹)	Cost Overrun (%)	Estimated Days	Actual Days	Delay (days)
Mean	257.92	536,451.20	574,610.00	6.76	40.92	44.72	3.80
Median	275.00	566,409.00	590,640.00	6.65	42.00	43.00	4.00
Minimum	61	83,846	87,516	3.85	14	17	0
Maximum	462	839,393	916,279	9.75	85	92	7
Standard Deviation	117.47	201080.72	217376.68	2.06	18.43	19.16	2.31

Observations:

On average, **projects cost about ₹5.7 lakhs**, with the actual cost being ~7% higher than the quoted price.

The **average project size is ~258 sq. ft.**, but projects vary widely, from as small as 61 sq. ft. to as large as 462 sq. ft.

Projects are usually **planned for ~42 days but end up taking ~45 days**, meaning most projects are delayed.

The **average delay is 4 days**, with some projects going up to **7 days late**.

Cost overruns are consistent, averaging 6–7%, which shows estimates are usually too low.

Larger projects (more days) tend to have **bigger absolute delays**, confirming the correlation seen in earlier analysis.

4 Detailed Explanation of Analysis process/method

The main aim of this project was to understand why cost estimates and project timelines in contract-based furniture projects are often inaccurate, and then design simple solutions to improve them. To achieve this, a structured analysis process was followed. The steps included data collection, data preparation, statistical analysis, root cause identification, and tool development. In addition, Python was used for statistical analysis and visualization, making the process faster and more accurate.

4.1 Data Collection

The first step was to gather real project data. Primary data was manually collected from The Furniture Master Interiors over a six-month period (January 2025 – June 2025). Sources included quotations, purchase records, bills, site reports, and team feedback. For each project, key details such as quoted cost, actual cost, estimated days, actual days, project size, material choices, and key issues faced were recorded. In total, 25 project records were prepared for analysis.

4.2 Data Cleaning and Organization

Once the dataset was ready, it was cleaned and organized. Here, Python libraries such as pandas were used to load the Excel file and check for missing or inconsistent values. In the dataset, text fields such as “Key Issues” were grouped into clear categories like supplier delays, client changes, labor shortages, design errors, and material shortages. At the same

time, numeric fields were standardized into consistent units. The “Key Issues” column, which originally had entries like “supplier delay,” “client requested extra cabinets,” or “workers absent on site,” was carefully organized into categories such as supplier issues, client modifications, or labor shortages.This process made the dataset clean, reliable, and structured, ensuring that the information could be studied systematically and comparisons across projects became accurate and meaningful.

4.3 Descriptive Statistics and Summary

With the dataset cleaned, descriptive statistics were calculated. Python was then used to generate descriptive statistics such as mean, median, minimum, maximum, and standard deviation. For example:The average cost overrun was **6.76%**. The average delay was **3.8 days**. Project sizes ranged from **61 sq. ft. to 462 sq. ft.** These results were obtained using simple `df.describe()` commands in pandas.

These numbers helped to quantify the scale of the problem. Graphs such as histograms and scatter plots were also created to visualize trends. For example, a histogram of delays showed that most projects were delayed between **2–5 days**, while a scatter plot of quoted vs. actual cost showed that almost every project went over budget.

4.4 Correlation and Pattern Analysis

To find deeper patterns, correlations between variables were studied. Python’s NumPy and pandas functions were used to calculate correlations. The strongest finding was the correlation of **0.92** between Actual Days and Delay. This means that the longer a project runs, the more delay it tends to have. Another pattern observed was that client-requested changes were one of the leading causes of both cost overruns and delays. These insights gave clear direction for building practical solutions.

4.5 Root Cause Identification

The “Key Issues” column was studied in detail, and problems were grouped into four major categories: **Material delays** (late delivery from suppliers), **Labor shortages** (especially during festivals), **Client changes** (extra cabinets, design modifications) and **Design revisions** (mid-project changes that required rework). This step helped identify what was truly driving the problems behind inaccurate quotations and missed timelines.

4.6 Tool Development

Based on the findings, three tools were designed to help the company in the future:

A **Cost Estimation Template** to generate more accurate quotations using historical data and buffer margins.

A **Material Requirement Planner** to prepare a bill of materials early and avoid supplier delays.

A **Labor Scheduling Tracker** to map tasks and workforce availability more effectively.

4.7 Validation and Recommendations

Finally, the tools and insights were validated by comparing them with past results. Recommendations included adding proportional time buffers for large projects, creating stronger supplier agreements, and building a more flexible labor pool.

5 Results and Findings

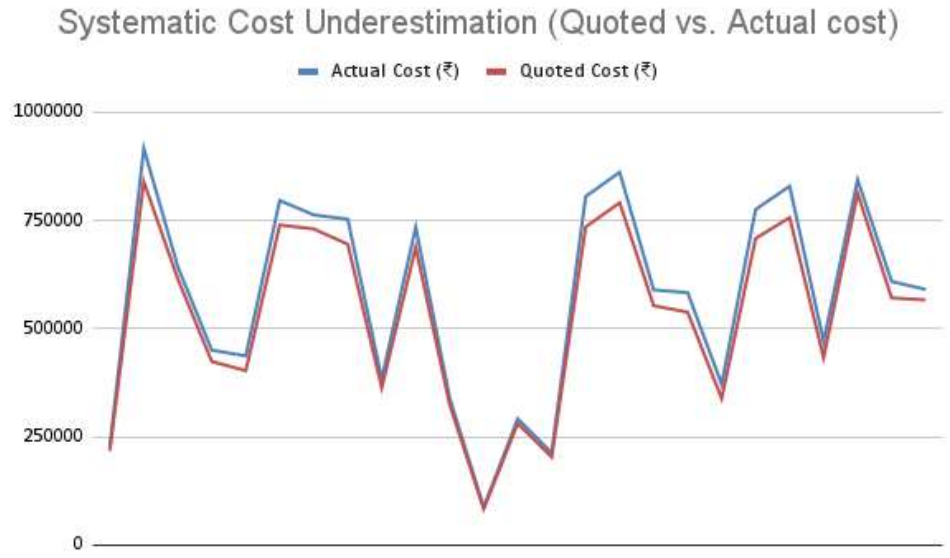
The analysis of The Furniture Master Interiors data using both **Excel-based visualization** and **Python modeling** revealed several key patterns and insights related to Quoted, Actual Cost, Actual Days and Delay Days etc.

5.1 Why are cost estimations at the quotation stage often inaccurate?

When we looked at the data from 25 projects, we saw a clear pattern: almost every project ended up costing more than what was originally quoted. On average, projects had a **6.76% cost overrun**, meaning the final spending was always higher than expected. That difference directly cuts into the company's profits.

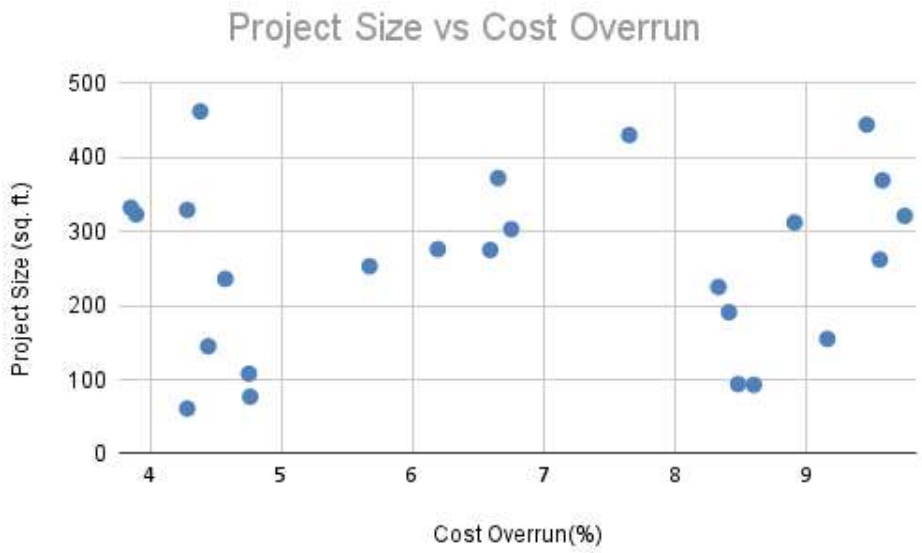
Relationship between the Actual Cost vs. Quoted cost.

(The relationship between Actual Cost and Quoted Cost is the most critical one in the dataset, as it directly defines profitability.)



The relationship between Quoted Cost and Actual Cost is perfectly positive but biased. Actual costs move directly with quoted costs, but with a consistent, positive error margin.

Relationship between the Project Size vs. Cost Overrun.



The relationship is defined by the formula for **Cost Overrun (%)**:

Actual Cost = Quoted Cost × (1 + (Cost Overrun% / 100))

Small projects (roughly under 150 sq.ft) – 6 projects in our data. These averaged about **5.9%** cost overrun (range ~4.3–8.6%). Most small jobs had overruns in the 4–6% range, with one notable exception up near 8.6%. Overall variation is moderate.

Medium projects (151–300 sq.ft.) – 8 projects. Average overrun was about **7.3%** (range ~4.6–9.6%). This group saw the highest average overrun, including two of the 9.5–9.8% cases. Medium jobs thus appear slightly more prone to larger overruns.

Large projects (>300 sq.ft.) – 11 projects. These averaged about **6.8%** overrun (range ~3.9–9.8%). The largest jobs tended toward slightly lower overruns on average than the medium ones, although one large job had nearly 9.5% overrun. The largest job (462 sq.ft.) had only 4.4% overrun.

Predictable and Significant Financial Impact

Finding: The magnitude of the difference is both significant and predictable.

Evidence: The average overrun is 6.76%. This means that on average, for every ₹100,000 quoted, the actual cost is ₹106,760.

Minimum Overrun: 3.85% (Project P023)

Maximum Overrun: 9.75% (Project P019)

Implication: The company can predict that it will lose, on average, 6.76% of its projected profit on every single project due to this underestimation.

Recommended Action: Revise the costing model to include a default risk buffer based on the historical overrun of **6.76%**, and adjust it for project-specific risks like project type, material choice, and client change history. Use size-aware contingency bands when quoting: **5–7% for small**, **8–10% for medium**, and **6–8% for large** projects.

To reduce surprises, lock down designs and finishes before work starts so scope is clear. Put a formal change-order process in place with clear pricing and approval steps for any client requests after the quote. Negotiate firm lead times or penalty clauses with key suppliers to cut material delays. Build a flexible labor plan — keep a backup pool or short-term hires for busy seasons.

These steps make quotes more realistic, protect profit margins, and cut mid-project firefighting. They also help set clearer expectations with clients and improve on-time delivery.

5.2 Why do project timelines regularly get delayed?

The data shows a very clear and strong relationship between the total time a project takes and the delay it experiences.

By definition, **Delay = Actual Days – Estimated Days**.

This means delay is directly tied to how long a project actually runs.

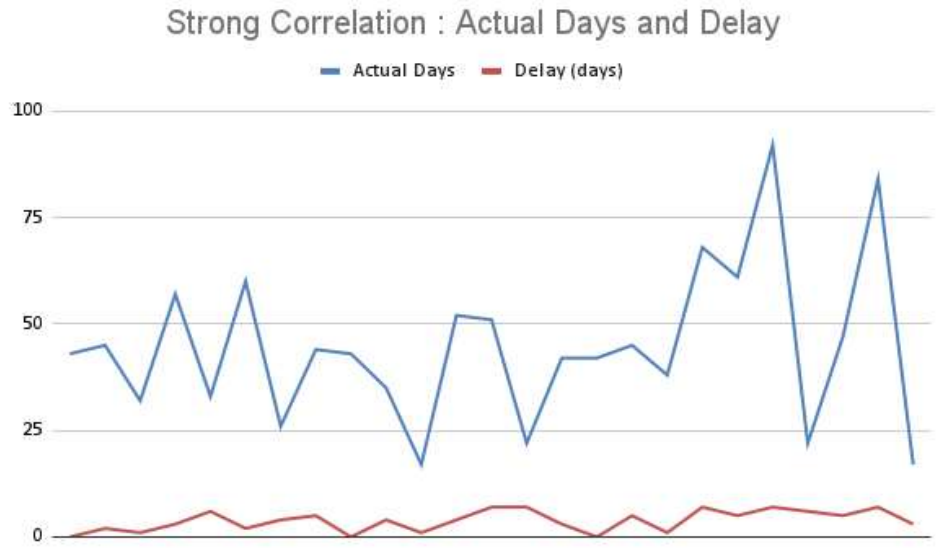
i. Extremely Strong Positive Correlation

Finding: There is a very strong, positive linear relationship between the total duration of a project and the length of its delay. Longer projects tend to have larger delays in terms of absolute days.

Evidence: A statistical analysis of the data reveals a Pearson correlation coefficient of approximately 0.92 between Actual Days and Delay. This value is very close to +1, indicating an almost perfect positive linear relationship.

Interpretation: This means that the number of Actual Days is the single biggest predictor of the absolute delay. As the actual duration increases, the delay increases in a very predictable way.

ii. Visual Evidence of the Relationship



We can also see this visually in the scatter plot with a trendline. The line slopes upwards to the right, showing that as the number of Actual Days increases, the Delay also increases in a very predictable way. On average, projects are delayed by **3.8 days**, but some of the larger projects had delays of up to **7 days**.

Conclusion and Business Implication

Currently, the company adds a fixed number of buffer days (for example, 2 days) for every project. While this might help in small projects, it is completely inadequate for larger ones. A 2-day buffer doesn't mean much when projects can easily be delayed by a week or more.

Recommended Action: The company should adopt a proportional buffering strategy, where buffer time is a percentage of the estimated duration. This ensures that larger projects get larger buffers, making timelines more realistic and reducing client dissatisfaction.