Winter Internship Project Work On

Training and Development

At

Central Coalfields Limited (CCL), Ranchi

Quality Control Department

By

KAUSHIK VATS

BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI

ACKNOWLEDGEMENT

I would like to express my sincere and heartfelt gratitude to **Central Coalfields Limited (CCL), Ranchi**, for providing me with the valuable opportunity to undergo my **Winter Internship Training** in the **Quality Control Department**. This industrial training has been a significant learning experience and has helped me gain practical exposure to quality assurance practices followed in the coal industry.

I am deeply indebted to the **Departmental Head and the respected officers and technical staff of the Quality Control Department**, Central Coalfields Limited, for their constant guidance, encouragement, and technical support throughout the training period. Their willingness to share knowledge, explain operational procedures, and involve me in day-to-day laboratory and field activities greatly enhanced my understanding of coal quality analysis and control mechanisms.

I would also like to extend my sincere thanks to the **Training and Human Resource Department, Central Coalfields Limited**, for organizing the training programme and for providing a disciplined, professional, and supportive work environment during the entire duration of the internship.

I take this opportunity to express my profound gratitude to my **faculty coordinator and the Department of Chemical Engineering, BIT Mesra**, for granting permission and providing academic guidance to undertake this industrial training as part of my curriculum.

Company Profile of Central Coalfields Limited (CCL), Ranchi

Central Coalfields Limited (CCL) is a subsidiary of Coal India Limited (CIL) and a Government of India undertaking under the Ministry of Coal. Headquartered at Ranchi, Jharkhand, CCL plays a vital role in meeting the coal requirements of the country, particularly catering to the energy needs of thermal power plants, steel plants, cement industries, and other core sectors.

CCL primarily operates in the coal-rich regions of Jharkhand, covering major coalfields such as North Karanpura, South Karanpura, Ramgarh, Giridih, Bokaro, and Hazaribagh. The company is engaged in both opencast and underground mining operations, with a greater emphasis on high-capacity opencast mines to enhance productivity and efficiency. CCL produces mainly non-coking coal, which is a critical input for power generation and industrial processes.

The organization follows a structured operational framework encompassing exploration, mining, coal beneficiation, quality control, transportation, and dispatch. Quality assurance is a key focus area, and CCL has well-equipped Quality Control laboratories to ensure compliance with coal grade specifications as per Indian standards. Modern techniques such as mechanized mining, conveyor systems, rapid loading systems, and advanced laboratory instruments are increasingly being adopted to improve operational performance.

Central Coalfields Limited also emphasizes sustainable mining practices, environmental protection, and corporate social responsibility (CSR). Initiatives related to land reclamation, afforestation, water management, and community development are actively undertaken. With a skilled workforce, strong technical expertise, and a commitment to national energy security, CCL continues to contribute significantly to India’s industrial and economic growth.

SWOT Analysis of Central Coalfields Limited (CCL), Ranchi

Strengths

Central Coalfields Limited is a subsidiary of Coal India Limited, which provides strong financial stability, policy support, and operational reliability. The company has access to abundant coal reserves in Jharkhand and benefits from well-established mining infrastructure and an experienced workforce. Its structured quality control systems ensure consistent coal grading and compliance with industrial requirements.

Weaknesses

CCL faces challenges related to aging mining equipment and partial dependence on conventional mining practices, which can impact productivity. As a public sector undertaking, procedural complexities and slower decision-making processes may affect operational flexibility. Environmental and land acquisition issues also pose constraints in certain operational areas.

Opportunities

Growing domestic demand for coal, especially from power and core industries, offers sustained market opportunities. Modernization through advanced mining technologies, automation, and data-driven quality control can significantly enhance efficiency and safety. Government initiatives focused on energy security further support long-term growth.

Threats

Stringent environmental regulations, increasing focus on renewable energy, and competition from alternative fuels pose long-term challenges. Social and environmental concerns related to mining activities may affect expansion plans and operational continuity.

Operational Risk & Decision Analytics

Prepared by: Kaushik Vats

Role: Intern

Organization: Central Coalfields Limited (CCL)
Duration: (Internship Period)

Tools Used: Excel, Python, Power BI, Regression

Background

Central Coalfields Limited (CCL) manages coal dispatch operations through both rail and road modes across multiple mining areas and source units. Each dispatch record contains critical operational and quality-related information such as mode of dispatch, area, source unit, month, calorific value (GCV), ash percentage, and notified coal price.

Traditionally, these large datasets are reviewed using spreadsheets and static reports, which limits the ability to quickly identify operational risks, quality deviations, pricing inconsistencies, and data integrity issues. As operations scale, management requires **data-driven decision support systems** that provide early warning signals, highlight high-risk patterns, and enable informed corrective actions.

This project aims to convert raw operational and pricing data into **actionable insights** using analytics techniques such as dashboards, trend analysis, and regression modelling. A concise management dashboard and structured analytical approach were developed to improve operational transparency and risk awareness.

.Goals

Principal goals of the original scope (brief and quantifiable):

Create a dashboard that is easy to use for management that summarizes important operational and quality threats.

Use a straightforward measure to quantify risk: Risk equals Likelihood × Frequency (with impact proxy).

Use regression analysis to examine the price-quality relationship (GCV → Price) and measure sensitivity.

Determine the underlying reasons by source unit and area, then emphasize the areas of concentration.

Provide a pilot implementation path and a quantifiable mitigation strategy.

Objectives

The primary objectives of this project are:

- To analyse rail and road dispatch operational data and identify key operational risks.
- To quantify risk using an explainable framework based on **likelihood and frequency**.
- To examine the relationship between coal quality (GCV) and notified prices using regression analysis.
- To identify root causes of high-risk patterns across areas, source units, and dispatch modes.
- To design a management-friendly dashboard that supports monitoring and decision-making.
- To propose a practical risk mitigation and implementation plan aligned with operational realities.

3. Data Description & Key Observations

The dataset used in this study consists of dispatch records spanning multiple operational years and includes more than **24,000 samples**. Key variables analysed include dispatch mode (rail/road), area, source unit, month, GCV, ash percentage, and price.

Key Observations:

DispatchModeConcentration:

Rail dispatch (SHRA) dominates operations, contributing a significantly higher number of records compared to road dispatch (SHRO). This indicates higher operational exposure and risk concentration in rail logistics.

- **SourceUnitSkewness:**
  A small number of source units (notably MAGADH OCP and AMRAPALI OCP) contribute a disproportionately large share of total records, increasing dependency risk.
- **DataQualityGaps:**
  A notable percentage of records lack proper source-unit identification, which reduces traceability and limits effective root-cause analysis.
- **SeasonalTrends:**
  Monthly analysis reveals peak operational volumes during specific months, indicating the need for proactive planning and resource allocation.
- **Outliers:** A limited number of records show unusually high price or ash deviations, contributing disproportionately to overall operational risk.

4. Methodology

Data Preparation

- Consolidated multiple Excel datasets into a unified structure.
- Standardised field names and formats for consistency.
- Converted GCV bands into numeric mid-values for analysis.
- Identified and flagged missing or inconsistent entries.
- Removed exact duplicates to ensure data accuracy.

Exploratory Analysis

- Computed KPIs such as total samples, mode-wise counts, and missing data percentages.
- Analysed frequency distributions of price differences and ash percentage.
- Studied monthly trends to identify operational peaks.

Risk Quantification

Risk was conceptualised as:

Risk ≈ Likelihood × Frequency × Impact

- **Frequency:** How often an issue occurs.
- **Likelihood:** Probability of recurrence across time or locations.
- **Impact:** Magnitude of price deviation, ash penalty, or operational disruption.

Risk scores were categorised into **Low, Medium, and High** for prioritisation.

5. Regression Analysis: Price vs GCV

Purpose

To quantify the relationship between coal quality (GCV) and notified price, and to understand the rationale behind incremental price adjustments.

Approach

- Used GCV mid-values for grades G1 to G17.
- Applied linear regression with:
  - **Independent variable (X):** GCV (kcal/kg)
  - **Dependent variable (Y):** Price (₹/tonne)

Interpretation

- The regression shows a **strong positive linear relationship** between GCV and price.
- The slope represents the price increase per unit increase in GCV.
- A high R^2 value indicates that coal quality explains most of the price variation.

Pricing Logic Explanation

Coal prices are notified by GCV bands. Within each band, **pro-rata adjustments** are applied based on actual calorific value. This explains non-rounded prices such as ₹3570 or ₹3420, which are systematic and policy-driven rather than arbitrary.

<u>CODE FOR REGRESSION</u>

```
1)import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# Manually creating the data (G1 to G17)
data = {
"Grade":
["G1","G2","G3","G4","G5","G6","G7","G8","G9","G10","G11","G12","G13","G14","G15","G16","G17"],

# GCV midpoint values (from CCL table)
"GCV_Mid": [7100, 6850, 6550, 6250, 5950, 5650, 5350, 5050, 4750,
```

```
4450, 4150, 3850, 3550, 3250, 2950, 2650, 2350],

# Price to Power (Rs/tonne)
"Price": [3570, 3420, 3260, 2980, 2520, 2100, 1918, 1600, 1250,
1130, 975, 906, 837, 768, 610, 524, 467]
}

df = pd.DataFrame(data)
df
```

2)

```
X = df[['GCV_Mid']]
y = df['Price']

model = LinearRegression()
model.fit(X, y)

print("Intercept (β0):", model.intercept_)
print("Slope (β1):", model.coef_[0])
print("R² value:", model.score(X, y))

plt.figure(figsize=(8,5))
plt.scatter(X, y, color='blue', label='Actual Prices')
plt.plot(X, model.predict(X), color='red', label='Regression Line')
plt.xlabel("GCV (kcal/kg)")
plt.ylabel("Price (Rs/tonne)")
plt.title("Coal Price vs GCV Regression (CCL Data)")
plt.legend()
plt.grid(True)
plt.show()
```

Explanation of Regression Analysis

- Coal quality, represented by **mid-point calorific value (GCV_Mid)**, was selected as the independent variable, as coal pricing is primarily based on calorific value.
- Coal **price (Rs/tonne)** was selected as the dependent variable to analyse how pricing varies with changes in coal quality.
- A **linear regression model** was applied to quantify the relationship between coal quality and price using the equation:

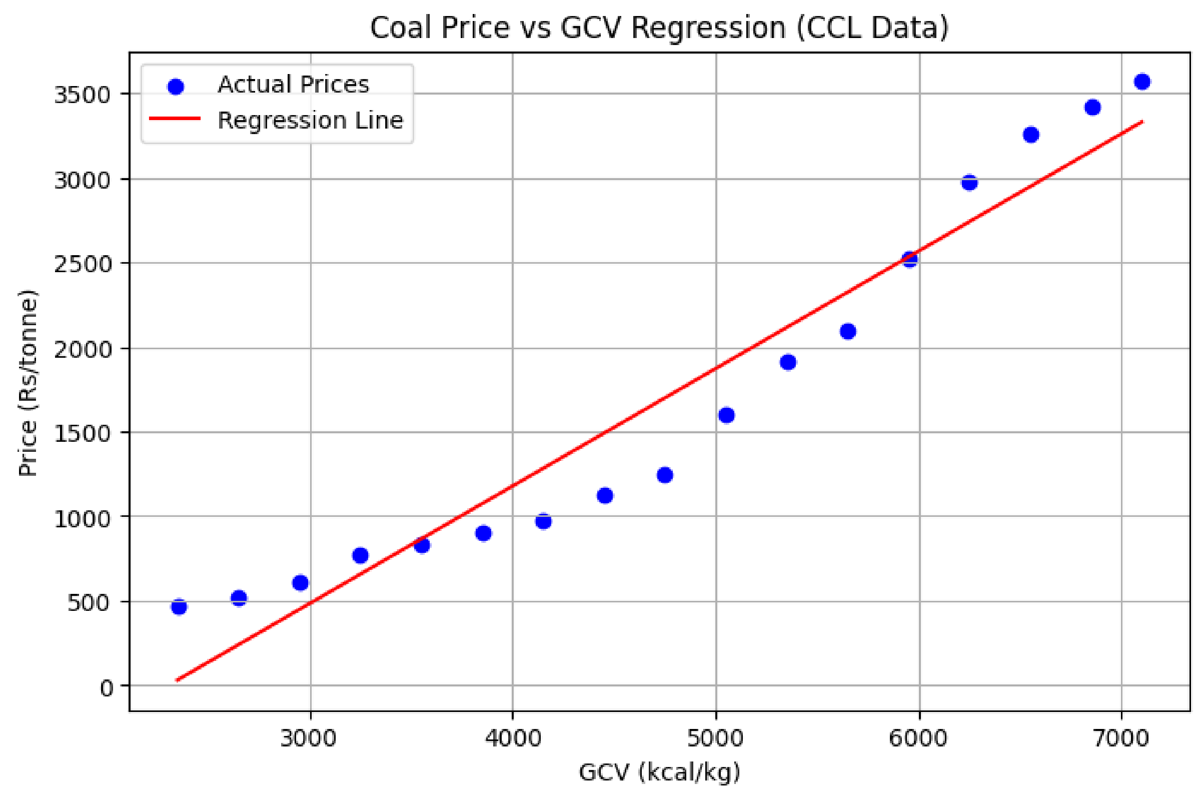
$$Price = \beta_0 + \beta_1 \times GCV$$

- The **intercept (β_0)** represents a reference price level, while the **slope (β_1)**

indicates the increase in price corresponding to a unit increase in GCV.

- The **coefficient of determination (R^2)** was used to assess the strength of the relationship and to measure how much of the price variation is explained by coal quality.
- The scatter plot represents actual price observations across different GCV levels, providing a clear view of data distribution.
- The regression line represents the best-fit trend, confirming a **positive and near-linear relationship** between GCV and coal price.
- The observed incremental price differences within the same grade are explained by **pro-rata calorific value adjustments**, where prices vary proportionally with actual GCV values inside a grade band.
- Overall, the regression analysis validates that coal pricing follows a structured, quality-linked mechanism rather than arbitrary adjustments.

GRAPH OF THE PLOT



Intercept (β0): -1598.9792566913904
Slope (β1): 0.6941468074814577
R^2 value: 0.9347762566167221

- THE SLOPE ($\beta_1 = 0.694$) TELL US

For **100 kcal/kg increase in GCV ----->** **Price increases by ~₹69 per tonne**
The regression slope shows that price increases by approximately ₹70 for every 100 kcal increase in GCV, which matches the pro-rata pricing mechanism used in notified coal prices.

- $R^2 = 0.935$ TELL US

93.5% of the variation in coal price is explained by GCV alone.
Interpretation:

- Coal quality is the **dominant factor** in pricing.
- Only ~6.5% of price variation comes from:

Logistics
Policy adjustments
Operational factors
The high R^2 value shows that coal quality explains most of the price variation, indicating a strong and consistent quality-based pricing system.

6. Dashboard Analysis & Graph Interpretation

Mode Distribution Chart
Shows rail dominance, highlighting where monitoring and capacity planning should be prioritised.

Area & Source Unit Distribution
Reveals balanced area representation but skewed source-unit dependence, signalling diversification risk.

Monthly Trend Chart
Identifies operational peaks, supporting proactive planning and staffing decisions.

Outlier Charts (Price & Ash)
Highlight a small set of records that contribute disproportionately to risk, enabling targeted corrective action.

Each chart is designed to provide **one insight and one recommended action**, keeping dashboards decision-focused

7. Comprehensive Risk Mitigation Plan

1. Standardise GCV measurement and lab calibration.
2. Enforce mandatory fields for critical data points.
3. Mode-wise monitoring with contingency planning.
4. Seasonal capacity planning for peak months.
5. Automated outlier detection and review process.
6. Validation of pro-rata pricing adjustments.
7. Weekly risk dashboard dissemination.
8. Periodic audits of pricing and quality data.
9. Staff training and accountability ownership.
10. Pilot-based implementation followed by phased scale-up

8. Objectivity in the Plan

The proposed risk mitigation and decision analytics framework is designed to be **objective, measurable, and data-driven**, thereby reducing dependence on subjective judgement and individual experience. Objectivity in this plan is ensured through the use of **quantifiable Key Performance Indicators (KPIs)** and statistically derived insights rather than assumptions or anecdotal observations.

Key parameters such as **data completeness**, **frequency of deviations**, and **price–quality regression fit (R^2 value)** form the foundation of decision-making. For example, data completeness can be objectively measured as the percentage of records containing mandatory fields such as source unit, GCV, and price. Similarly, deviation frequency is calculated based on the number of occurrences where price or quality parameters exceed predefined thresholds, enabling consistent identification of high-risk events.

The inclusion of regression analysis further strengthens objectivity by mathematically validating the relationship between coal quality and pricing. The regression slope and R^2 value provide numerical evidence of how strongly price is driven by GCV, removing ambiguity from pricing interpretation. Changes in these values over time can also be tracked to assess improvements or emerging risks.

By defining clear thresholds, numerical benchmarks, and repeatable analytical methods, the plan ensures transparency and auditability. Management decisions derived from this framework can be justified using data and documented metrics, thereby improving governance, accountability, and long-term operational consistency.

9. Effectiveness Measurement

The effectiveness of implementation will be evaluated using clearly defined and measurable performance indicators. One key metric is the reduction in missing or incomplete data records, which directly reflects improvement in data governance. A consistent decline in such gaps indicates successful enforcement of mandatory data capture rules.

Another critical measure is the decrease in unresolved outliers, particularly in price and quality parameters. Tracking the number of flagged cases and their resolution time helps assess the responsiveness and efficiency of the risk management process. Faster closure of high-impact deviations signifies improved operational control.

The price–quality alignment, measured through changes in the regression R^2 value over time, provides a quantitative indicator of pricing consistency. An improving or stable high R^2 suggests better adherence to quality-linked pricing principles.

Collectively, these metrics enable continuous monitoring of system performance, ensuring that the implemented framework delivers tangible risk reduction, operational transparency, and sustained decision support.

10. Benefits to Coal & FMCG Supply Chains

- Improved predictability of coal quality.
- Reduced pricing disputes.
- Early detection of quality risks.
- Enhanced confidence in supplier–buyer relationships.
- Better planning for downstream industries.

11. Conclusion

This project successfully demonstrates how operational and quality-related data can be converted into meaningful, decision-ready insights through the application of analytics. By integrating operational dashboards with statistical analysis, the study provides a structured approach to identifying risks, understanding operational patterns, and supporting management decisions in coal dispatch operations across rail and road modes.

The dashboard analysis offered a consolidated view of dispatch volumes, mode dependence, source-unit concentration, monthly trends, and high-impact outliers. These visual insights enabled quick identification of operational risk areas, data quality gaps, and periods of peak activity that require proactive planning. The use of simple, management-friendly visualisations ensured that complex operational data could be interpreted easily and acted upon effectively.

The regression analysis further strengthened the analytical depth of the project by quantitatively validating the relationship between coal quality (GCV) and notified price. The strong positive slope and high coefficient of determination confirmed that coal pricing follows a systematic, quality-linked mechanism based on pro-rata calorific value adjustments. This analysis also explained observed incremental price differences within grade bands, reinforcing transparency and consistency in pricing practices.

Overall, the project demonstrates a scalable and objective framework for operational risk monitoring and decision support. The proposed methodology, dashboards, and mitigation measures enhance data governance, improve pricing clarity, and enable proactive risk management. When implemented in a phased and systematic manner, this approach can significantly improve operational resilience, transparency, and governance across CCL operations.