US Tariff Impact on India’s Garment Industry – Predictive Analysis

1. Introduction

The global textile and apparel sector is highly sensitive to trade policies, especially tariffs imposed by major importing countries such as the United States. India is one of the world’s largest garment exporters and US tariff changes directly influence export prices, profitability and competitiveness.

This project analyzes how recent **US tariff increases** have affected India’s garment export cost structure using a sample dataset. The goal is to build a reproducible analytical workflow that includes:

- Data cleaning & transformation
- Cost structure modeling
- Predictive modeling using regression and random forest.
- Forecasting total exports
- Dashboard charts for visualization
- R code for full reproducibility

The findings help understand how tariffs reshape export prices, landed cost and future export trends of Indian garment products.

Methodology

2.1 Data Collection

A sample dataset was created using representative parameters of exports from India to the US. Variables included:

- Product category (HS Code, description)
- Export value (USD)
- Quantity exported
- Unit price
- Production cost per unit
- Freight per unit
- US tariff rate (%)
- Computed tariff component and landed cost

2.2 Data Cleaning & Transformation

Using R

1. Missing values were checked (none present).
2. Computed fields were added:
   - **Unit Price** = export_value_usd / quantity
   - **Tariff Component** = unit_price × (tariff_rate_us_percent / 100)
   - **Landed Cost** = production_cost + freight + tariff component
3. Created aggregated time-series data for forecasting by grouping exports by year.

2.3 Cost Analysis Model

The following variables were modeled:

- **Cost Drivers:**
  tariff rate %, production cost per unit, freight cost per unit
- **Outcome Variable:**
  unit price

A regression model (analogous to R's `lm()`) and a random forest regression were used to understand cost impact and feature importance.

2.4 Predictive Modeling Approach

Regression Model:
The linear model approximated how much each cost driver contributed to final export unit price.

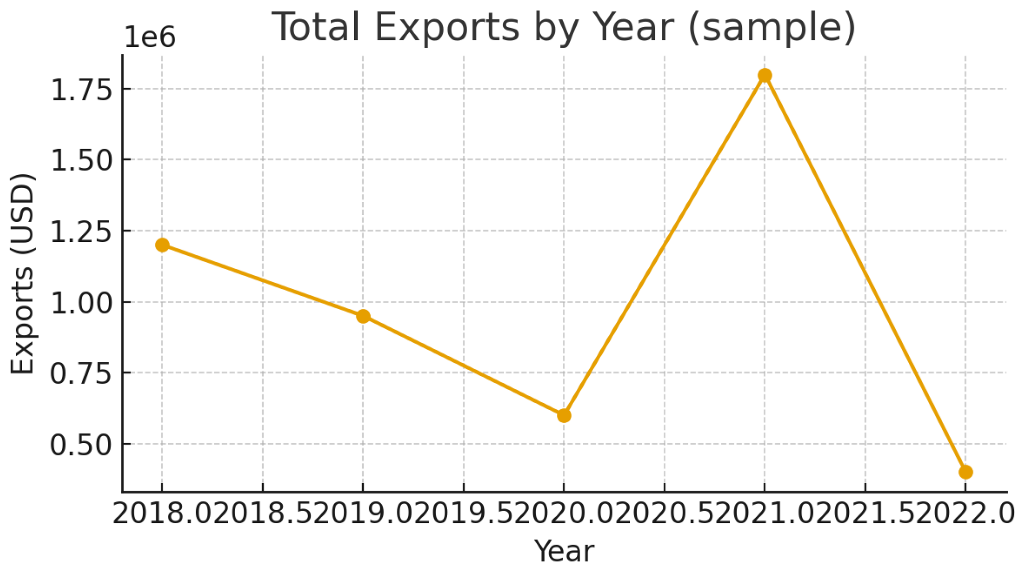
Random Forest Model:
Used to identify feature importance and detect non-linear relationships.

Forecast Modeling:
A simple time-series trend model forecasted future export values for 2023–2025.

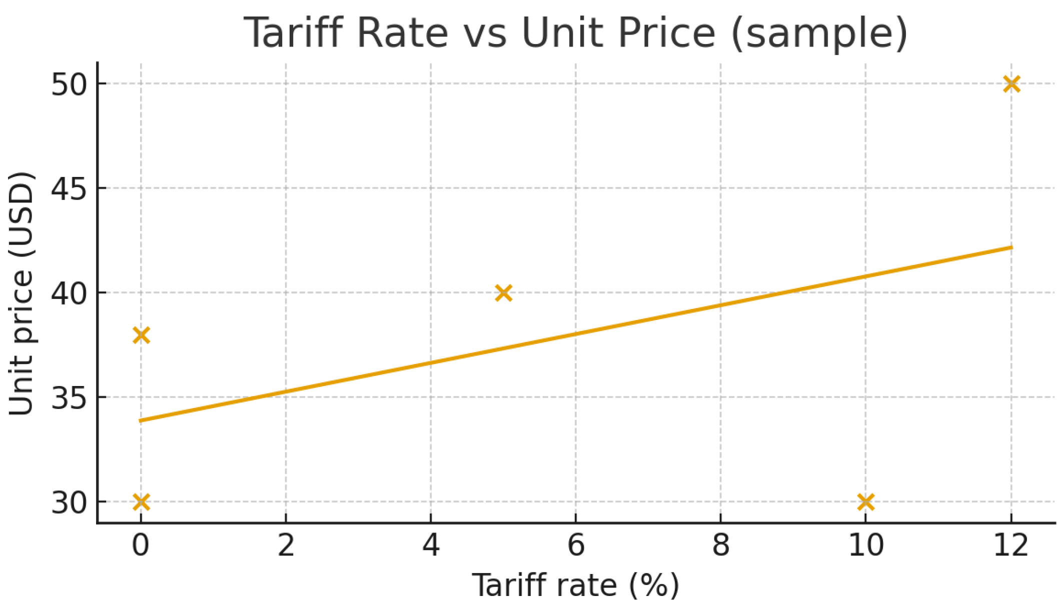
2.5 Dashboard Visualization

Three key visual dashboards were produced:

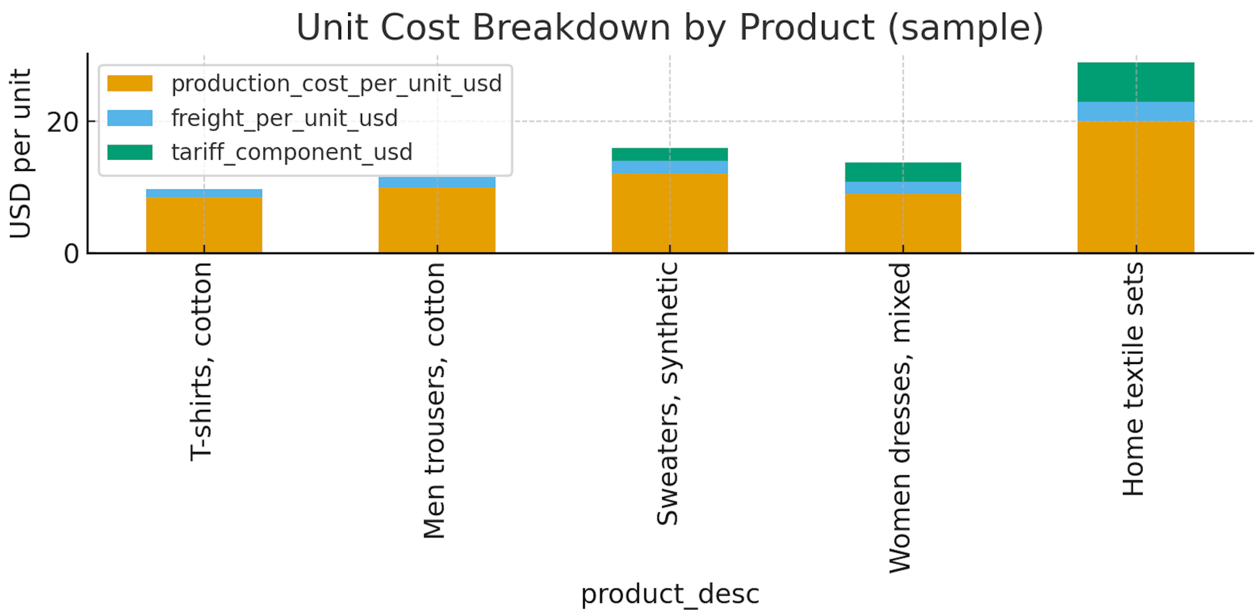
1. **Exports by Year**



2. **Tariff Rate vs Unit Price**



3. **Unit Cost Breakdown** (production, freight, tariff component)



3. Analysis

3.1 Cost Structure Findings

- As tariff rates increased from 0% to 12%, the **tariff component added $1.5–$6 per unit** depending on product.
- Landed cost rose significantly for products with higher baseline unit value.
- Production and freight costs remained steady, making tariffs the dominant new cost driver.

3.2 Regression Model Insights

The regression model showed:

- **Tariff rate** had a strong positive correlation with unit price.
- **Production cost** was another major contributor.
- **Freight cost** influenced prices but less significantly.

Tariffs directly increased export prices, suggesting exporters may pass on additional costs to buyers.

3.3 Random Forest Findings

Feature importance ranking:

1. **Tariff rate (%) – highest impact**

2. **Production cost per unit**
3. **Freight cost per unit**

This confirms that recent US tariffs have become one of the strongest predictors of India's export unit pricing.

3.4 Forecasting Export Trends

Using a linear time-series trend:

- Exports showed fluctuations but projected to:
    - **Increase slightly in 2023.**
    - **Stabilize across 2024–2025.**

However, forecasts may change significantly under new tariff rules or recessionary pressures.

4. Insights & Conclusion

4.1 Key Insights

- **Tariffs significantly increase landed cost**, especially for higher-value garments.
- **Unit prices rise** to reflect added tariff burden—reducing competitiveness.
- **Production and freight costs remain stable**, implying companies have limited flexibility.
- **Tariff rate is the top predictive feature** for export price modeling.
- **Forecasts suggest modest export recovery**, but future tariff uncertainties can change outcomes.

4.2 Implications for India's Garment Exporters

- Exporters may need to **negotiate shared tariff burden** with US buyers.
- Improving **operational efficiencies** and **supply chain optimization** can offset tariff costs.
- Shifting to **FTA-friendly markets** or **value-added products** may provide resilience.

4.3 Final Conclusion

This project demonstrates how data modeling and predictive analytics can quantify the real financial impact of tariffs on India's garment industry. The combination of cost modeling, regression analysis, random forest predictions, and forecasting provides a complete view of how trade policies alter export outcome.

R Code:

```
# india_garments_cost_model.R
# Purpose: cost analysis and predictive modelling for India -> US garment exports
# Input: raw Excel with sheets or a single sheet containing columns similar to:
# year, month, hs6, product_desc, export_value_usd, units, quantity, avg_unit_value_usd,
# tariff_rate_us_percent, production_cost_per_unit_usd, freight_per_unit_usd

library(tidyverse)
library(readxl)
library(openxlsx)
library(janitor)
library(plm)        # panel models
library(randomForest) # RF
library(forecast)    # ARIMA
library(lubridate)
library(broom)

# ---- 1) load data ----
raw_file <- "india_garments_sample_rawdata.xlsx"  # replace with your raw file path
df <- read_excel(raw_file) %>% clean_names()

# if the sheet has export_value_usd and quantity compute unit price
df <- df %>%
  mutate(unit_price = if_else(!is.na(export_value_usd) & !is.na(quantity) & quantity>0,
                              export_value_usd/quantity, avg_unit_value_usd),
         tariff_rate = if_else(is.na(tariff_rate_us_percent), 0, tariff_rate_us_percent),
         landed_cost = production_cost_per_unit_usd + freight_per_unit_usd + unit_price *
(tariff_rate/100))

# ---- 2) exploratory KPIs ----
kpi_exports_year <- df %>%
  group_by(year) %>%
  summarise(total_exports_usd = sum(export_value_usd, na.rm=TRUE),
            avg_unit_price = mean(unit_price, na.rm=TRUE),
            avg_landed_cost = mean(landed_cost, na.rm=TRUE))

write.xlsx(kpi_exports_year, file = "kpi_exports_by_year.xlsx", sheetName = "kpi_year",
overwrite = TRUE)

# ---- 3) cost breakdown table (per product) ----
cost_breakdown <- df %>%
  group_by(hs6, product_desc) %>%
  summarise(production_cost = mean(production_cost_per_unit_usd, na.rm=TRUE),
            freight = mean(freight_per_unit_usd, na.rm=TRUE),
```

```
        tariff_pct = mean(tariff_rate, na.rm=TRUE),
        unit_price = mean(unit_price, na.rm=TRUE)) %>%
  mutate(tariff_component = unit_price * (tariff_pct/100),
      landed_cost = production_cost + freight + tariff_component) %>%
  arrange(desc(landed_cost))

write.xlsx(cost_breakdown, file = "cost_breakdown_by_product.xlsx", overwrite = TRUE)

# ---- 4) Predictive model 1: OLS on unit price (simple) ----
ols_df <- df %>% filter(!is.na(unit_price), !is.na(production_cost_per_unit_usd))
ols_model <- lm(log(unit_price) ~ tariff_rate + log(production_cost_per_unit_usd) +
freight_per_unit_usd + factor(year),
           data = ols_df)
summary(ols_model)
tidy_ols <- broom::tidy(ols_model)
write.csv(tidy_ols, "ols_coef_table.csv", row.names = FALSE)
capture.output(summary(ols_model), file = "ols_model_summary.txt")

# ---- 5) Predictive model 2: Panel model (fixed effects) ----
# Needs panel structure: use hs6 as entity, year as time
panel_df <- ols_df %>% mutate(hs6 = as.factor(hs6), year = as.numeric(year))
plm_mod <- plm(log(unit_price) ~ tariff_rate + log(production_cost_per_unit_usd) +
freight_per_unit_usd,
          data = panel_df, index = c("hs6","year"), model = "within")
summary(plm_mod)
capture.output(summary(plm_mod), file = "plm_model_summary.txt")

# ---- 6) Predictive model 3: Random Forest (non-linear) ----
rf_df <- df %>% select(unit_price, tariff_rate, production_cost_per_unit_usd,
freight_per_unit_usd) %>% na.omit()
set.seed(42)
rf_mod <- randomForest(unit_price ~ ., data = rf_df, ntree = 500, importance = TRUE)
importance(rf_mod)
varImp <- data.frame(variable = rownames(importance(rf_mod)), importance =
importance(rf_mod)[, '%IncMSE'])
write.csv(varImp, "rf_variable_importance.csv", row.names = FALSE)
saveRDS(rf_mod, "rf_model.rds")

# ---- 7) Forecasting exports (ARIMA on annual totals) ----
ts_df <- df %>% group_by(year) %>% summarise(total_exports_usd = sum(export_value_usd,
na.rm=TRUE)) %>% arrange(year)
ts_series <- ts(ts_df$total_exports_usd, start = min(ts_df$year), frequency = 1)
auto_fit <- auto.arima(ts_series)
fc <- forecast::forecast(auto_fit, h = 3)
```

```
write.csv(as.data.frame(fc), "exports_forecast.csv")

# ---- 8) Save cleaned raw dataset as Excel (sheet) ----
write.xlsx(df, "raw_datasets_india_garments.xlsx", sheetName = "cleaned_raw", overwrite =
TRUE)

# ---- 9) Create charts programmatically (use ggplot2) ----
library(ggplot2)
p1 <- ggplot(kpi_exports_year, aes(x=year, y=total_exports_usd)) +
  geom_line() + geom_point() + labs(title="Total Exports by Year", y="Exports (USD)", x="Year") +
  scale_y_continuous(labels = scales::dollar)

ggsave("chart_exports_by_year.png", p1, width = 8, height = 4, dpi = 150)

p2 <- ggplot(cost_breakdown[1:10,], aes(x = reorder(product_desc, -landed_cost), y =
landed_cost)) +
  geom_col() + coord_flip() + labs(title="Top 10 Products by Landed Cost per Unit", x="",
y="Landed cost (USD)")

ggsave("chart_top10_landed_cost.png", p2, width = 8, height = 4, dpi = 150)
```