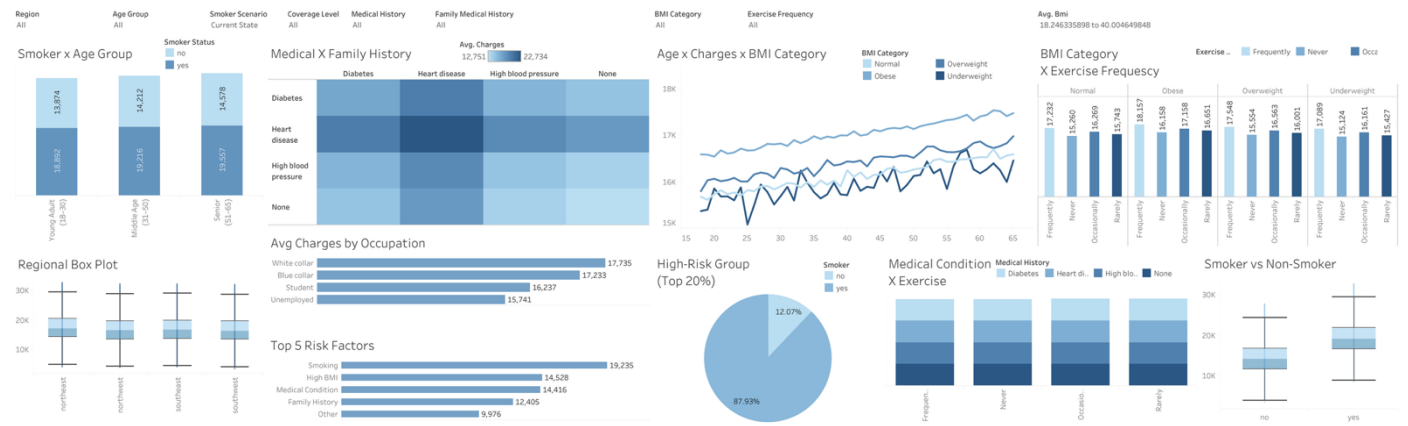


# Implementation

## Section 1: The Dashboard



The insurance dashboard's main goal is to help them make informed decisions. The dashboard combines data on costs, risks, and health, providing a reliable foundation for planning, pricing, and resource allocation throughout the insurance industry. It highlights which risk factors have the biggest impact, such as smoking or high BMI. It also tells exactly who is incurring the highest claims, letting you target policies or interventions where they're most effective. It is designed for people who need to analyse the numbers underwriters, healthcare providers, and government officials. They rely on it to understand how various factors contribute to rising healthcare costs. It helps to transform a large dataset of policyholder information into strategic intelligence that supports cost control, risk mitigation, and public health improvement across the entire system.

## Section 2: The Dataset

### 1. Dataset Overview and Purpose:

The dataset used for this dashboard contains 1,000,000 individual insurance records, offering a robust foundation for statistical analysis and risk modelling. Its scale ensures that insights remain stable even after applying multiple filters, such as age groups, BMI categories, medical histories, or regional splits. Because the data mirrors real-world insurance profiles, the patterns uncovered align closely with what underwriting teams, healthcare analysts, and policy agencies typically observe. The size and diversity of the dataset also minimize sampling bias, allowing the dashboard to reveal meaningful cost trends and behaviour-driven risk differences.

### 2. Attribute Composition and Analytical Value:

The data consists of 12 structured attributes, covering demographics (age, gender, region), lifestyle variables (smoking status, exercise frequency, BMI), financial measures (charges, coverage level), and health indicators (medical history and family medical history). Numerical attributes such as age, BMI, and charges enable continuous trend analysis, while categorical dimensions support segmentation through bar charts, heatmaps, box plots, and distribution-based visuals. Many attributes replicate fields commonly found in insurance intake forms, making the dataset suitable for realistic modelling of costs, risk scoring, and eligibility decisions. Importantly, each attribute has a clear analytical purpose. For example, smoking and BMI serve as modifiable lifestyle risks, while medical histories reflect long-term chronic disease exposure.

### 3. Data Quality, Pre-processing, and Feature Engineering:

A major advantage of this dataset is its perfect completeness, with zero missing values across all 12 attributes. This eliminated the need for traditional cleaning steps such as imputation, outlier removal, or type corrections, allowing the focus to shift toward meaningful pre-processing and feature engineering within Tableau. Several transformations were applied to enhance analytical clarity. For example, BMI

was binned into medically relevant categories (Normal, Overweight, Obese), and age was grouped into life-stage segments (Young Adult, Middle Age, Senior) to support comparative visuals.

Additional calculated fields were created to derive combined risk indicators (e.g., smoker + high BMI), normalize occupation groups for consistent segmentation, and generate interaction-based fields such as medical vs. family history combinations for the heatmap.

Tableau-level enhancements included percentile-based filtering to isolate top-cost segments, coverage-tier grouping, and standardized formatting of categorical values for consistent dashboard behaviour. Because the underlying data was already clean and structured, these engineered fields significantly enriched the analytical depth that allowed the dashboard to reveal nuanced patterns, compare multi-factor risk relationships, and maintain smooth performance even during heavy interactivity.

### **Section 3: Dashboard Users**

#### **1. Insurance Underwriting Team**

The underwriting team relies on the dashboard to understand how key risk factors like age, BMI, smoking, and medical history drive up insurance costs. With this information, they can assess risk accurately, whether it's for an individual or an entire group. They also adjust premiums based on this data, making sure pricing reflects differences in lifestyle, location, and current market trends. By spotting patterns across different jobs and regions, they can see where pricing might need to be updated.

The dashboard also helps them quickly identify high-risk customers. Underwriters can then design tailored coverage or reach out with preventive measures. This makes their portfolio more stable and helps prevent unexpected claims.

#### **2. Healthcare Providers / Hospitals**

Hospitals and healthcare providers use the dashboard to get a sense of community health especially how habits like smoking, inactivity, or dangerous work conditions connect to chronic diseases. It lets them pinpoint health problems showing up in specific areas. They use these insights to design preventive care, launching wellness programs that match the behaviours most likely to cause illness. By addressing root causes instead of just symptoms, they can intervene sooner and lower long-term expenses.

On the operational side, they look at demographic and cost trends to plan staffing, equipment, and facilities. This keeps them ready for patient demand and helps them use resources efficiently.

#### **3. Government Health Agencies**

Government health agencies use the dashboard to monitor trends like obesity and smoking, tracking how much these issues impact the healthcare system. This data guides their decisions and shapes long-term policy. They also check the dashboard to see whether health programs like subsidies or awareness campaigns are really working across different regions or groups. With this feedback, they can hold programs accountable and decide where to focus future resources.

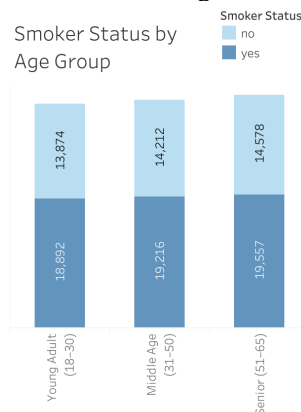
Additionally, agencies examine differences in medical costs between regions and social groups. By connecting health outcomes to things like occupation or income, they can design policies that improve access and reduce inequality.

## Section 4: Questions

1. How does smoking status impact insurance charges across different age groups?
2. What is the cost gap between smokers and non-smokers?
3. Is there a correlation between BMI categories and insurance costs and how does exercise frequency moderate this relationship?
4. What are the top 5 risk factors contributing to high insurance costs?
5. How do average insurance charges vary across different regions and what factors drive regional disparities?
6. Which occupations have the highest average insurance costs?
7. What is the demographic profile of the highest risk group?
8. How does the combination of medical history and family medical history influence insurance charges across coverage levels?
9. What is the distribution of medical conditions across the population and how do they vary by occupation and exercise frequency?
10. How do insurance charges progress with age, and what role do modifiable factors play in cost escalation?

## Section 5: Plots

1. How does smoking status impact insurance charges across different age groups?

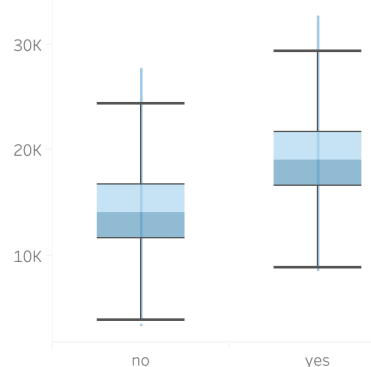


### Stacked/Grouped Bar Chart

This chart shows how smoking habits shift as people age. The proportion of smokers and non-smokers in each age group can be seen side by side. The stacked bars make it obvious when smoking peaks in life. That helps for spotting age-specific risk patterns that influence insurance charges.

2. What is the cost gap between smokers and non-smokers?

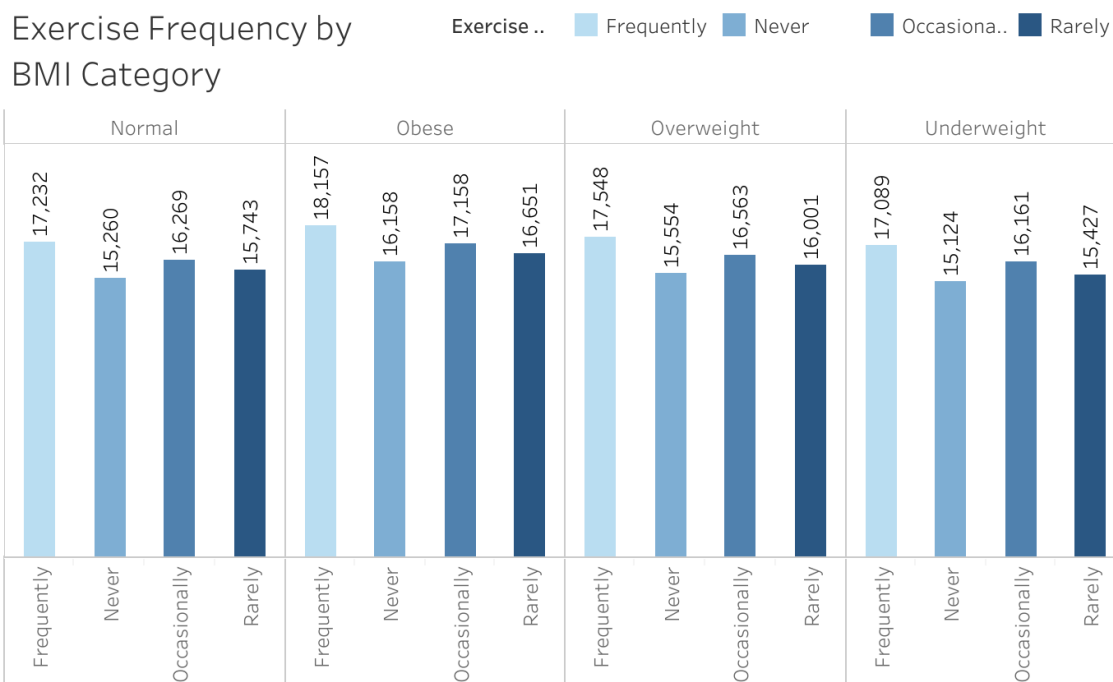
Smoker vs Non-Smoker  
Charges Distribution



### Box Plot

The chart lays out the real cost difference between smokers and non-smokers. It allows us to compare the full range of charges for each group, including the medians and any outliers. The shape clearly shows that smokers consistently incur higher charges, quantifying the financial penalty of smoking.

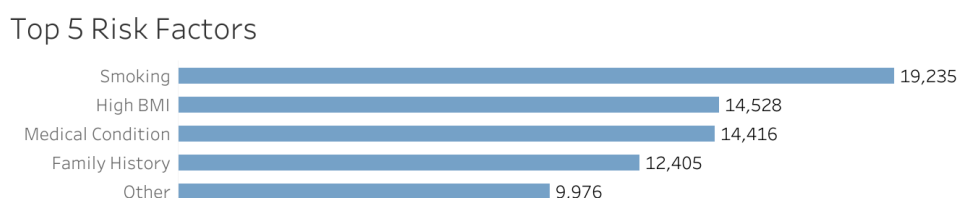
3. Is there a correlation between BMI categories and insurance costs, and how does exercise frequency moderate this relationship?



#### Clustered Bar Chart

The chart breaks down how the weight and exercise habits work together to affect insurance costs. Bars grouped by BMI show the average charges, while each group splits further by exercise frequency. It's easy to observe if being active helps lower expenses, even in higher-risk categories like obesity. It's a straightforward way to see the real impact of an active lifestyle.

4. What are the top 5 risk factors contributing to high insurance costs?

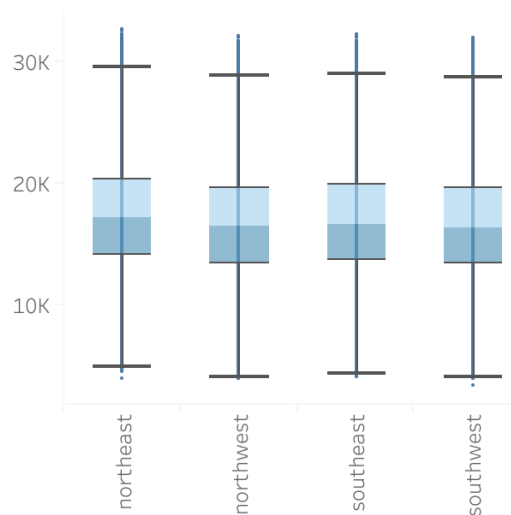


#### Ranked Bar Chart

This plot identifies the leading contributors of what really drives up insurance costs—smoking, high BMI, age, chronic conditions, and so on. Each bar's length shows how much each factor matters. It's an easy way to see what insurers worry about most and where to focus if they want to cut costs.

5. How do average insurance charges vary across different regions and what factors drive regional disparities?

### Charges by Region



#### Box Plot

This chart compares average insurance charges across different regions by comparing medians, spreads, and high-cost extremes. Regional differences stand out, revealing how local health habits, healthcare expenses, or demographic population shape the numbers. It displays how some regions just end up carrying a heavier financial load.

#### 6. Which occupations have the highest average insurance costs?

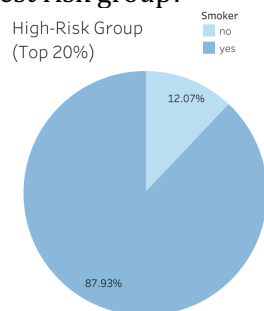
### Avg Charges by Occupation



#### Sorted Bar Chart

The chart sorts jobs from the lowest to highest average insurance costs. It's easy to see which professions pay the most and which pay the least. Patterns start to emerge and maybe certain work environments or stress levels push costs up. It's a gives an insight on which groups might need extra wellness support.

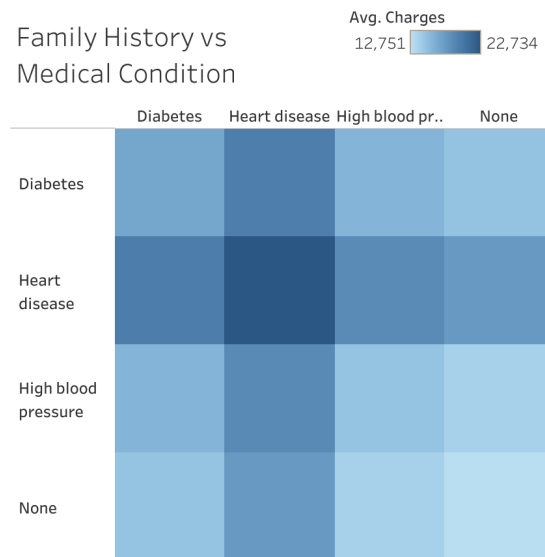
#### 7. What is the demographic profile of the highest risk group?



#### Donut Chart

The chart focuses on the top 20% of insurance claims as the real heavy hitters. It breaks down who's in this expensive group, whether by smoking status or other demographics. We can immediately tell which group takes up the biggest slice. That's useful for targeting risk-mitigation strategies where they matter most.

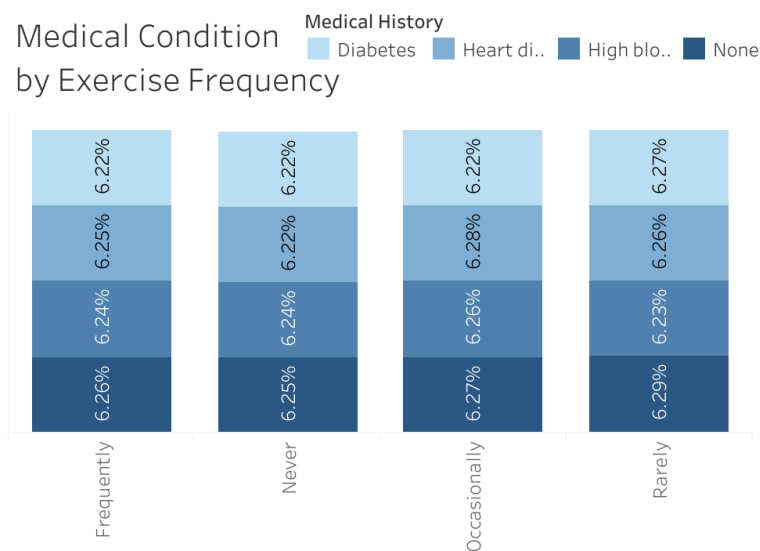
#### 8. How does the combination of medical history and family medical history influence insurance charges across coverage levels?



### Heatmap

The heatmap shows how personal medical history combined with your family's history affects the charges. The darkest cells point to the "toxic combinations" where risks really compounded to pile up. It clarifies which hereditary and personal health mixes drive premium escalation across coverage levels.

9. What is the distribution of medical conditions across the population and how do they vary by occupation and exercise frequency?

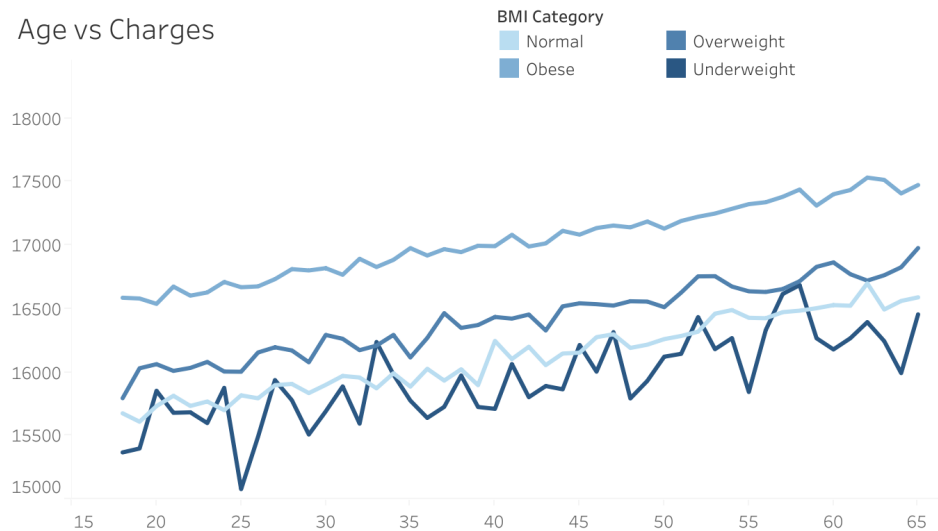


### 100% Stacked Bar Chart

The chart tracks how medical conditions spread out across people with different activity levels. Each bar sums up the condition mix for a specific exercise group, so we can compare active people to those who are less active. It displays which conditions are most influenced by lifestyle choices.

10. How do insurance charges progress with age, and what role do modifiable factors play in cost escalation?

Age vs Charges



### Multiple Line Chart

The chart addresses how age and weight together climb cost over time. There are separate lines for people with different BMI category of weight. As the years add up, the cost gap widens, showing how much BMI can speed up age-related increases in insurance charges.

## Section 6: Interactivity

### 1. Age Group Filter

This filter allows users to switch between Young Adult, Middle Age, and Senior groups and updates the charts on smoking status, cost trends, and risk comparisons right away. It identifies age-related patterns across different visuals without digging around.

### 2. Smoker Scenario Filter

The control toggles to switch between smokers and non-smokers. The grouped bar chart, smoker vs non-smoker box plot, and cost comparisons all update instantly. It enables users to instantly view how smoking affects costs and changes the data.

### 3. BMI Category Filter

Selecting a different BMI group, updates BMI × Exercise bar chart and age–cost trend lines. This helps users to see how weight categories shift risk and cost patterns.

### 4. Exercise Frequency Filter

This filter adjusts the chart to see how activity level changes impact the BMI × Exercise cost chart and the stacked bars showing medical conditions. It gives a clearer view of how exercise shapes costs and health profiles.

### 5. Region Selector

The filter helps to change regions, and the regional cost box plot along with any geography-based summaries gets updated. It allows users to compare healthcare costs by location and spot regional differences fast.

### 6. Coverage Level Filter

This control updates the medical × family history heatmap, filtering average charges for basic, standard, or premium plans. It clarifies how risk combinations behave under different insurance tiers.

### 7. Medical History Condition Selector (for Family and Self)

This selector updates the distribution chart by breaking it down across exercise or occupation groups. This lets us dig into health conditions and how they play out in different demographics.

8. Tooltip Drill-Down (Hover Interactivity)

Hovering over any bar, line, or box for details like exact charges, counts, and percentages. It provides more info instantly, without crowding the main chart.

9. Avg BMI Range Selector

The BMI range slider lets users narrow the analysis to individuals within a specific BMI interval, dynamically updating charts such as BMI × Exercise, Age × Charges × BMI lines, and overall cost distributions. This interactivity helps uncover how insurance charges and risk patterns shift when focusing on narrower weight profiles.

## Links

### Tableau Dashboard -

[https://public.tableau.com/views/ProjectDV-Group15\\_17648488173190/Dashboard-Group15?:language=en-US&publish=yes&:sid=&:redirect=auth&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/ProjectDV-Group15_17648488173190/Dashboard-Group15?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link)

### Mural Dashboard -

<https://app.mural.co/t/group15dv8892/m/group15dv8892/1763500714125/4dc1aaa82f36983bc3e737de2ae9015bdf3eb7a0?sender=ueae1aa4802689266eefd8638>