**REPORT ON INSURANCE DASHBOARD**

Introduction**:**

The Insurance dataset from Kaggle contains information about individuals and their insurance policies, including various features such as age, sex, BMI, number of children, smoking status, region, and insurance charges. This dataset can be used to explore the relationships between these features and insurance charges, and to build predictive models that can accurately predict insurance charges based on the available features.

Main Features:

The insurance dataset contains 1338 observations and 7 features, including:

Age: Age of the individual

Sex: Gender of the individual

BMI: Body mass index of the individual

Children: Number of children the individual has

Smoker: Smoking status of the individual

Region: Region the individual is from

Charges: Insurance charges for the individual

Main Goal:

The main goal of this dataset is to explore the relationships between the features and the insurance charges, and to build predictive models that can accurately predict insurance charges based on the available features. This can help insurance companies to better understand the factors that contribute to higher insurance charges and to develop more accurate pricing models.

**WH Questions**

Q: How does age affect insurance charges?

Based on the dataset, there is a positive correlation between age and insurance charges. This indicates that as individuals get older, their insurance charges tend to increase. For example, the average insurance charge for individuals over 60 years old is more than three times higher than the average insurance charge for individuals under 30 years old.

Q: Does smoking status affect insurance charges?

Yes, based on the dataset, individuals who smoke tend to have higher insurance charges compared to non-smokers. On average, smokers have insurance charges that are more than twice as high as non-smokers.

Q: Are there any differences in insurance charges based on region?

Yes, based on the dataset, individuals in different regions have different average insurance charges. For example, individuals in the southeast region tend to have higher insurance charges compared to individuals in other regions. The average insurance charge for individuals in the southeast is more than 25% higher than the average insurance charge for individuals in the southwest region.

Conclusion:

The insurance dataset is a valuable resource for exploring the relationships between different factors and insurance charges. By analyzing this dataset, we can gain insights into the factors that contribute to higher insurance charges, such as age, smoking status, and region. This can help insurance companies to develop more accurate pricing models and to better understand the risks associated with different groups of individuals. Additionally, this dataset can be used to build predictive models that can accurately predict insurance charges based on the available features, which can be used to help individuals make more informed decisions about their insurance options.

Other Insights:

In addition to the WHO questions, the insurance dataset can be used to explore other interesting questions, such as the relationship between BMI and insurance charges or the impact of having children on insurance charges. Furthermore, this dataset can be used to build a variety of predictive models, such as linear regression, decision trees, and neural networks, to accurately predict insurance charges based on the available features.