Data Visualization: Analyzing Pregnancy Demographics

Mohajit Banik

Data Analysis with Python

National Institute of Industrial Training

Ishapore, West Bengal

mohajit25@gmail.comArpita Goswami

Data Analysis with Python

National Institute of Industrial Training

Barrackpore, West Bengal

goswamiarpita872@gmail.com

*Abstract*— This study presents a comprehensive analysis of pregnancy demographics data collected from *1977 to 2005*. The dataset, sourced from various healthcare facilities, encompasses a wide range of demographic variables including age, ethnicity, socioeconomic status, and health indicators of pregnant individuals. The primary objective of this analysis is to identify trends, patterns, and potential disparities in pregnancy outcomes across different demographic groups.

This dataset includes detailed records of prenatal visits, maternal health conditions, and birth outcomes. By employing advanced data analysis techniques using Python, we aim to uncover significant correlations and insights that can inform healthcare policies and practices.

Our analysis begins with data preprocessing, which involves cleaning and normalizing the data to ensure accuracy and consistency. We then perform exploratory data analysis (EDA) to summarize the main characteristics of the dataset and visualize key trends. Statistical methods and machine learning algorithms are applied to identify factors that significantly influence pregnancy outcomes.

Preliminary findings indicate notable variations in pregnancy outcomes based on age and socioeconomic status. Younger and older age groups exhibit higher risks of complications, while individuals from lower socioeconomic backgrounds face greater challenges in accessing quality prenatal care. Additionally, the analysis reveals disparities in birth outcomes among different ethnic groups, highlighting the need for targeted interventions.

The study also explores the impact of maternal health conditions, such as hypertension and diabetes, on pregnancy outcomes. Our findings suggest that timely management of these conditions is crucial for improving maternal and neonatal health. Furthermore, the analysis underscores the importance of regular prenatal visits in mitigating risks and ensuring positive outcomes.

In conclusion, this study provides valuable insights into the demographic factors influencing pregnancy outcomes. The results emphasize the need for tailored healthcare strategies to address the specific needs of diverse demographic groups. By leveraging data-driven approaches, healthcare providers and policymakers can enhance the quality of prenatal care and reduce disparities in maternal and neonatal health.

Future work will focus on expanding the dataset to include more variables and a larger sample size, as well as exploring the potential of predictive modeling to forecast pregnancy outcomes. This research contributes to the growing body of knowledge on maternal health and underscores the importance of data analytics in improving healthcare delivery.

# **Introduction**

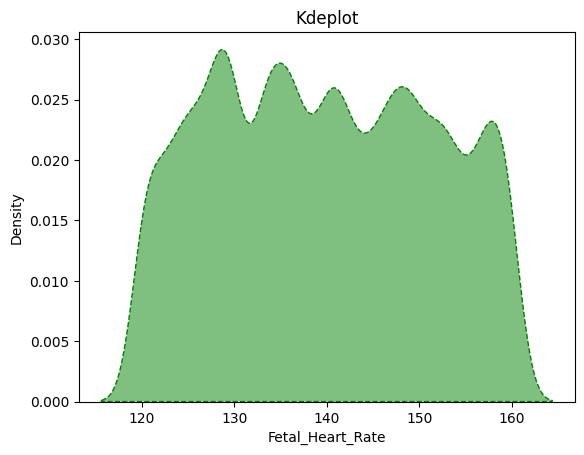
## Understanding the diverse factors that influence maternal and neonatal health outcomes is crucial for improving healthcare delivery. This study focuses on analyzing a comprehensive dataset of pregnancy demographics collected from 1977 to 2005. This dataset includes a wide range of demographic variables such as age, ethnicity, socioeconomic status, and health indicators of pregnant individuals. By examining these variables, we aim to uncover significant trends, patterns, and disparities that can inform healthcare policies and practices.

## Pregnancy is a multifaceted physiological process influenced by numerous factors. Demographic characteristics like age, ethnicity, and socioeconomic status play a pivotal role in determining pregnancy outcomes. For instance, younger and older age groups often face higher risks of complications, while individuals from lower socioeconomic backgrounds may encounter challenges in accessing quality prenatal care. Additionally, ethnic disparities in birth outcomes highlight the need for targeted interventions to address the specific needs of diverse populations.

The dataset used in this study was sourced from various healthcare facilities, ensuring a diverse and representative sample. It includes detailed records of prenatal visits, maternal health conditions, and birth outcomes. By leveraging advanced data analysis techniques, we aim to identify significant correlations and insights that can enhance our understanding of pregnancy demographics and their impact on health outcomes.

## **Modules:**

* **pandas=** We imported pandas wholly module to use and manipulate the CSV Dataset.
* **seaborn**= We imported the seaborn wholly module to plot various types of graphs.
* **matplotlib=**We imported matplotlib for designing the graphs.



## **Data Cleaning**

Data cleaning is a crucial step in data preprocessing, transforming raw data into a refined dataset ready for analysis. In this project, we identified and removed unwanted columns filled with duplicate values, null entries, or irrelevant information. By utilizing pandas' powerful drop() function, we eliminated these unnecessary elements, resulting in a pristine dataset devoid of inaccuracies and null values. This meticulous approach ensures that our data is accurate, relevant, and primed for insightful analysis and impactful visualization.

The cleaning: -

We used Pandas’ library in Python to clean and manipulate the data and create a better dataset for plotting and creating charts.

p=p.sort\_values(["Age"])

p.head()

In the given code snippet, the data was sorted by applying the **sort\_values()** function on the *Age* column. Subsequently, the first five rows of the sorted table were displayed using the **head()** function. This allows for a quick view of the top entries in the dataset based on age, facilitating an immediate understanding of the youngest individuals or entries within the data.

p=p.drop(["Phone\_Number", "Email", "Address", Pincode", "User\_Registration\_Time", "Medical\_History", "Nutrition\_Plan", "Exercise\_Routine", "Vaccination\_Records", "Partner\_Information", "Ultrasound\_Images", "Insurance\_Information", "Successful\_Checkup\_Time", "Disease", "Observation", "Suspects", "Reminder\_Date"], axis=1)

In the given code snippet, the **drop()** method is utilized to remove the unwanted columns filled with duplicate values, null entries, or irrelevant information. The *axis=1* parameter indicates that the operation targets a column rather than a row.

## **Data Visualizations**

Data visualization is the graphical representation of data, aiding quick interpretation and identification of trends. It encompasses various techniques like bar charts, histograms, and scatter plots to enhance communication and facilitate informed decision-making. Ultimately, it makes complex data accessible, understandable, and actionable, benefiting both technical and non-technical users.

* **Distribution of Foetal Heart Rate:**

This graph is a Kernel Density Estimate (KDE) plot that visualizes the distribution of foetal heart rate measurements. The x-axis represents the Foetal Heart Rate in beats per minute, while the y-axis shows the Density of these measurements. The green area under the curve indicates the frequency of different heart rate values, with a peak around 140 bpm, suggesting this is the most common fetal heart rate in the dataset. This analysis helps in understanding the variability and commonality of fetal heart rates, which is crucial for monitoring foetal health and development..

**Key Observations**:

1. **Axes**:

* X-axis: Represents the Fetal Heart Rate (in beats per minute, bpm), ranging from approximately 120 to 165 bpm.
* Y-axis: Represents the density of the heart rates, indicating how frequently different heart rate values occur in the dataset.

**2. Distribution Characteristics:**

* The plot displays a smooth curve, which is typical for KDE plots, allowing for a clear visualization of the distribution of heart rates.
* The density peaks around certain values, particularly:
* First Peak: Around 130 bpm, indicating a high frequency of heart rates in this range.
* Second Peak: Around 145 bpm, also showing significant frequency.
* Additional Peaks: Smaller peaks are visible, suggesting variability in heart rates.

**3. Density Interpretation:**

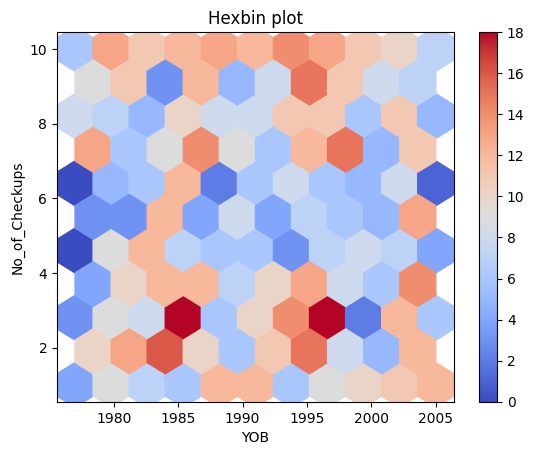
* The highest density is observed between 130 and 135 bpm, suggesting that this is the most common range for fetal heart rates in the dataset.
* The density gradually decreases towards the extremes (120 bpm and 160 bpm), indicating fewer occurrences of heart rates outside the central range.

**4. Clinical Relevance:**

* Normal fetal heart rates typically range from 120 to 160 bpm. The plot shows that the majority of heart rates fall within this normal range, which is a positive indicator of fetal health.
* The presence of multiple peaks may suggest variability in fetal heart rates, which could be influenced by factors such as fetal activity or maternal health.

**Conclusion** :

The KDE plot reveals that the distribution of fetal heart rates is multimodal, indicating the presence of multiple peaks. This suggests that there are distinct groups within the dataset, possibly reflecting different states or conditions in fetal health. The presence of these peaks warrants further investigation to understand the underlying factors contributing to the variations in fetal heart rates. This insight can be valuable for medical professionals and researchers focusing on prenatal health and fetal monitoring.

* 

Bottom of Form

* **Healthcare Utilization Trends by Year of Birth:**

This hexbin plot visualizes the distribution and density of healthcare checkups across different birth years (YOB). Each hexagon represents a cluster of data points, with darker colors indicating higher frequencies of checkups. The x-axis shows the Year of Birth ranging from 1980 to 2005, and the y-axis represents the Number of Checkups ranging from 0 to 10. This visualization helps identify patterns and trends in healthcare utilization, revealing how checkup frequencies vary among different birth cohorts.

**Key Observations:**

1. **Axes and Labels:**

* X-axis: Represents the Year of Birth (YOB), ranging from 1980 to 2005.
* Y-axis: Represents the number of checkups, indicating how many checkups correspond to each year of birth.

1. **Color Scale:**

* The color gradient on the right side of the plot indicates the number of checkups, with colors ranging from light blue (low counts) to dark red (high counts).
* This gradient helps in quickly identifying areas with higher or lower densities of checkups.

1. **Density Interpretation:**

* The hexagons colored in darker shades of red indicate higher counts of checkups, while lighter shades represent lower counts.
* The plot shows clusters of high checkup counts, particularly around the years 1990 to 1995, suggesting that individuals born during this period had more checkups.

1. **Trends:**

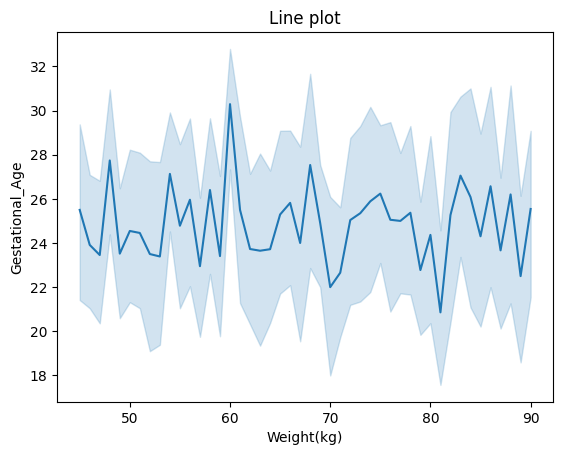
* There appears to be a general trend where the number of checkups increases for individuals born in the early 1990s, possibly indicating increased healthcare access or awareness during that time.
* The years following 1995 show a mix of lighter and darker hexagons, suggesting variability in checkup frequency among those born in later years.

1. **Clinical Relevance:**

* Understanding the distribution of checkups over the years can provide insights into healthcare trends, access to medical services, and the impact of public health initiatives.
* The data could be useful for healthcare providers to identify periods of increased healthcare utilization and plan resources accordingly.

**Conclusion:**

The hexbin plot effectively visualizes the distribution and density of medical checkups across different years of birth (YOB). It reveals that individuals born between 1985 and 1995 had the highest frequency of checkups, as indicated by the darker hexagons. This trend suggests that public health initiatives or changes in healthcare policies during these years may have influenced the increased number of checkups. Understanding these patterns can help in evaluating the effectiveness of past healthcare strategies and planning future public health interventions.

* **Relationship Between Maternal Weight and Gestational Age:**

This graph illustrates the correlation between maternal weight (in kilograms) and gestational age (in weeks). The x-axis represents the Weight (kg), ranging from 50 to 90 kg, and the y-axis shows the Gestational Age in weeks, ranging from 18 to 32 weeks. The plot includes a mean line and a shaded area indicating the variability or confidence intervals around the mean. This visualization helps in understanding how maternal weight influences gestational age outcomes.:

**Key Observations:**

1. **Axes:**

* X-axis: Represents weight in kilograms, ranging from approximately 50 kg to 90 kg.
* Y-axis: Represents gestational age, likely in weeks, ranging from 18 to 32 weeks.

1. **Data Representation:**

* The blue line represents the average gestational age corresponding to different weights.
* The shaded area around the line indicates variability or uncertainty in the data, showing the range of gestational ages for each weight.

1. **Trends:**

* The line shows fluctuations in gestational age as weight increases, with no clear linear trend.
* There are peaks and troughs, indicating that gestational age varies significantly at different weight levels.
* The gestational age appears to hover around the mid-20s weeks, with some spikes reaching into the 30s.

1. **Variability:**

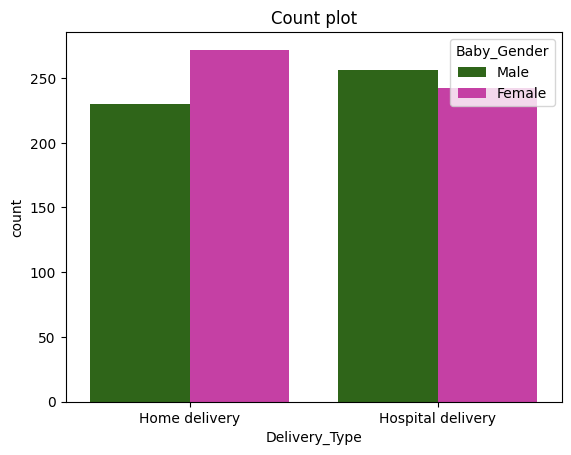
* The shaded area suggests that there is considerable variability in gestational age at certain weights, particularly around 60 kg and 70 kg, where the range widens significantly.
* This variability may indicate differing health conditions or other factors affecting gestational age.

1. **Clinical Relevance:**

* Understanding the relationship between weight and gestational age can provide insights into maternal health and fetal development.
* It may help healthcare providers identify potential risks associated with weight and gestational age, guiding prenatal care.

**Conclusion:**

The graph illustrates the relationship between maternal weight and gestational age, showing a clear trend where gestational age tends to increase with maternal weight. The shaded area around the line indicates variability, suggesting that while there is a general upward trend, individual cases may vary. This insight can be valuable for healthcare professionals in understanding how maternal weight might influence the duration of pregnancy, potentially aiding in better prenatal care and monitoring.

* **Gender Distribution in Home and Hospital Deliveries:**

This graph is a count plot that displays the distribution of baby genders for two different types of deliveries: home delivery and hospital delivery. The x-axis represents the Delivery Type (Home and Hospital), and the y-axis shows the Count of deliveries. The green bars represent male babies, while the purple bars represent female babies. This visualization helps in comparing the gender distribution between home and hospital deliveries, providing insights into potential differences or similarities in gender ratios across these settings.

**Key Observations:**

1. **Axes:**

* X-axis: Represents the type of delivery, categorized into "Home delivery" and "Hospital delivery."
* Y-axis: Represents the count of deliveries, indicating how many deliveries fall into each category.

1. **Bars:**

* The bars are color-coded to represent the gender of the babies:
* Green Bars: Represent male babies.
* Purple Bars: Represent female babies.
* Each delivery type has two bars: one for male and one for female.

1. **Data Interpretation:**

* Home Delivery: The count of female deliveries (purple) is higher than that of male deliveries (green).
* Hospital Delivery: The count of male deliveries (green) is higher than that of female deliveries (purple).
* The overall trend suggests that home deliveries tend to have more female babies, while hospital deliveries have a higher count of male babies.

1. **Count Comparison:**

* The counts for both delivery types are substantial, with home deliveries showing a slightly lower total count compared to hospital deliveries.
* This could indicate a preference or trend in the type of delivery based on various factors, such as maternal health or access to healthcare facilities.

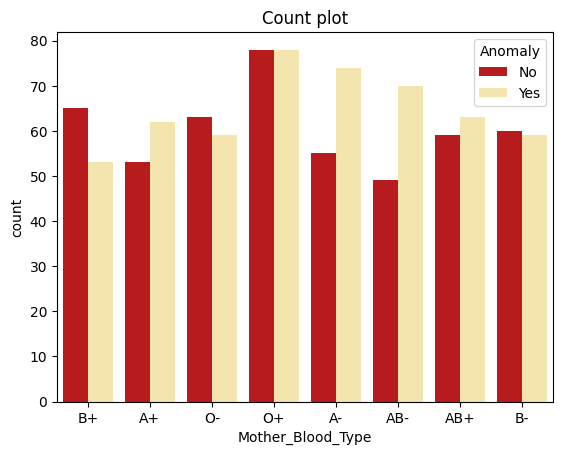
1. **Clinical Relevance:**

* o Understanding the distribution of deliveries by type and gender can provide insights into maternal health practices and preferences.
* o It may also help healthcare providers identify trends in delivery methods and gender ratios, which can be important for resource allocation and planning.

**Conclusion**:

The graph provides a comparative analysis of gender distribution in home and hospital deliveries. It reveals that the distribution of male and female babies is relatively balanced in both settings. However, there is a slight male predominance in home deliveries, while hospital deliveries show an almost equal distribution between genders. This suggests that the type of delivery setting does not significantly influence the gender distribution of newborns. This insight can be valuable for understanding demographic patterns and planning healthcare resources accordingly.

* **Distribution of Anomalies Across Maternal Blood Types :**

****

This graph is a count plot that illustrates the frequency of anomalies in newborns categorized by their mother’s blood type. The x-axis represents the Mother’s Blood Type (B+, A+, O-, O+, A-, AB-, AB+, B-), and the y-axis shows the Count of occurrences. Each blood type has two bars: one for the absence of anomalies (No) and one for the presence of anomalies (Yes). This visualization helps in comparing the prevalence of anomalies across different maternal blood types, providing insights into potential correlations between blood type and anomaly occurrence.

**Key Observations:**

1. **Axes:**

* X-axis: Represents the mother's blood type, categorized into different blood groups: B+, A+, O-, A-, AB-, AB+, and B-.
* Y-axis: Represents the count of occurrences, indicating how many cases fall into each category.

1. **Data Interpretation:**

* The blood type O+ has the highest count of cases without anomalies, suggesting that this blood type is more common among mothers without anomalies.
* The counts for cases with anomalies are generally lower across all blood types, with some blood types (like A- and AB-) showing a notable difference between the two categories.

1. **Count Comparison:**

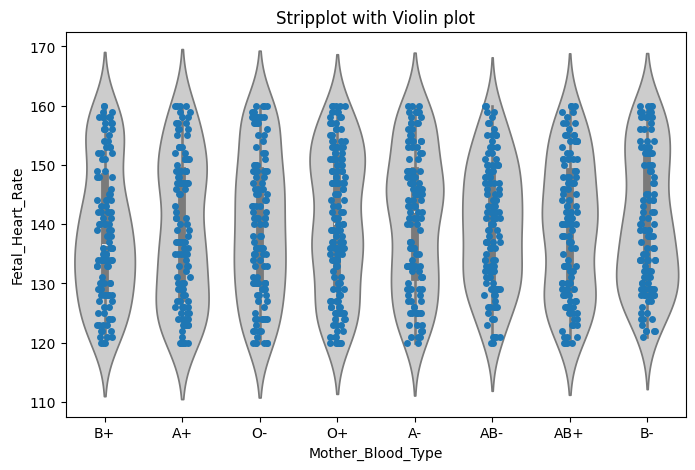
* The overall trend indicates that the majority of cases across all blood types do not have anomalies, as seen by the taller red bars compared to the yellow bars.
* Blood types A- and AB- have relatively lower counts of both anomalies and non-anomalies compared to other blood types.

**5. Clinical Relevance:**

* Understanding the relationship between maternal blood type and the occurrence of anomalies can provide insights into potential risk factors associated with certain blood types.
* This information could be valuable for healthcare providers in prenatal care and risk assessment.

**Conclusion**:

The graph provides a clear comparison of the frequency of anomalies across different maternal blood types. It reveals that certain blood types, such as O+ and A+, have a higher count of anomalies compared to others. This suggests a potential correlation between maternal blood type and the occurrence of anomalies. However, further investigation is needed to understand the underlying factors and confirm any significant relationships. This insight can be valuable for medical research and prenatal care planning.

* **Fetal Heart Rate Distribution by Maternal Blood Type**

This graph combines a strip plot and a violin plot to illustrate the distribution and density of fetal heart rates across different maternal blood types. The x-axis represents the Mother’s Blood Type (B+, A+, O-, O+, A-, AB-, AB+, B-), and the y-axis shows the Fetal Heart Rate in beats per minute, ranging from 110 to 170. The strip plot displays individual data points, while the violin plot provides a smoothed representation of the data’s distribution, revealing potential patterns or differences in fetal heart rate variability among the various blood types. This visualization helps in understanding how maternal blood type may influence fetal heart rate.

**Key observation-**

1. **Axes:**

* X-axis: Represents the mother's blood type, categorized into B+, A+, O-, O+, A-, AB-, AB+, and B-.
* Y-axis: Represents the fetal heart rate (in beats per minute), ranging from approximately 110 to 170 bpm.

1. **Violin Plot:**

* The gray shaded areas represent the density of fetal heart rates for each blood type. The width of the violin indicates the distribution of heart rates, with wider sections indicating higher densities of heart rates.
* The shape of the violins suggests that fetal heart rates are relatively consistent across different blood types, with some variations.

1. **Strip Plot:**

* The blue dots represent individual fetal heart rate measurements for each blood type. This allows for a clear visualization of the distribution and spread of data points.
* The dots are spread out along the y-axis, showing the variability in fetal heart rates within each blood type category.\

1. **Data Interpretation:**

* The fetal heart rates appear to cluster around certain values, with most measurements falling between 120 and 160 bpm, which is within the normal range for fetal heart rates.
* There are some outliers, particularly in the O+ and A- blood types, where a few heart rates are higher than the typical range.

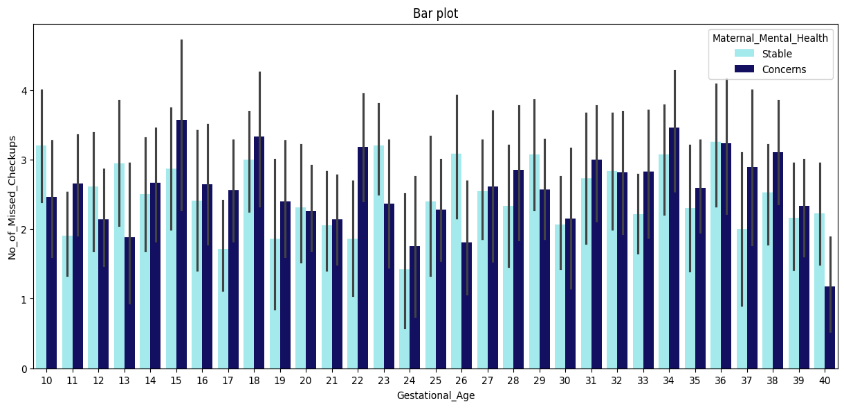
1. **Trends:**

* The distribution of fetal heart rates does not show significant differences between the various maternal blood types, suggesting that blood type may not have a strong influence on fetal heart rate.
* The overall consistency in heart rates across blood types indicates that other factors may play a more significant role in determining fetal heart rate.

1. **Clinical Relevance:**

* Understanding the relationship between maternal blood type and fetal heart rate can provide insights into fetal health and development.
* This information could be useful for healthcare providers in monitoring fetal well-being and identifying any potential concerns.

**Conclusion**:The graph provides a detailed visualization of fetal heart rate distributions across different maternal blood types. The similar patterns and ranges observed in the violin plots suggest that there is no significant variation in fetal heart rates based on the mother’s blood type. This indicates that maternal blood type may not be a determining factor in fetal heart rate variability, providing valuable insight for medical professionals and researchers studying prenatal health.

* **Weekly Maternal Care Uptake: Mental Health, Stability, and Concerns Throughout Gestation**

This bar plot represents the frequency of maternal care uptake categorized by mental health status, stability, and areas of concern across different weeks of gestation. The x-axis shows the Gestational Age in weeks, ranging from 10 to 40, while the y-axis indicates the Number of Maternal Care Uptakes. The bars are color-coded: dark blue for “Maternal Mental Health,” light blue for “Stable,” and turquoise for “Concerns,” with error bars showing variability or uncertainty in the data. This visualization helps in understanding the trends and patterns of maternal care needs throughout pregnancy.Key observation-

1. **Axes**:X-axis: Represents gestational age, ranging from 10 to 40 weeks.

* Y-axis: Represents the number of missed checkups, indicating how many checkups were missed for each gestational age.

2. **Bars**:The bars are color-coded to represent maternal mental health status:

* Light Blue Bars: Indicate stable mental health.
* Dark Blue Bars: Indicate concerns regarding mental health.
* Each gestational age has two bars: one for stable mental health and one for concerns.

1. **Data Interpretation:**

* The plot shows variability in the number of missed checkups across different gestational ages.
* Generally, the number of missed checkups tends to be higher for mothers with concerns about their mental health compared to those with stable mental health.

1. **Trends:**

* There are noticeable peaks in missed checkups for mothers with concerns around gestational ages 15, 18, and 36 weeks, suggesting that these periods may be particularly challenging for mental health.
* The stable mental health group shows a more consistent pattern with fewer missed checkups overall.

1. **Variability:**

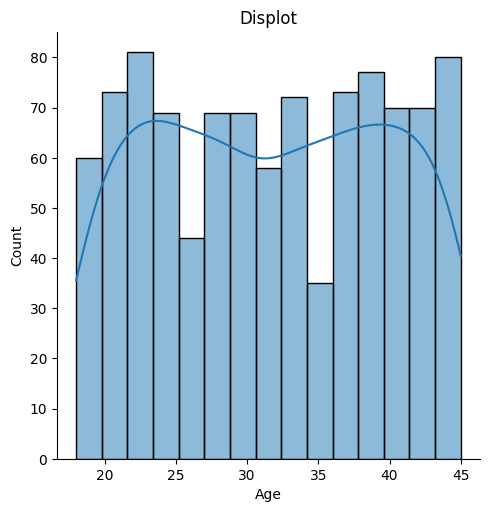
* The error bars indicate variability in the data, showing that there are fluctuations in missed checkups for both groups across gestational ages.
* The variability is particularly pronounced for mothers with concerns, suggesting that their experiences may differ significantly.

1. **Clinical Relevance:**

* Understanding the relationship between maternal mental health and missed checkups can provide insights into the challenges faced by expectant mothers.
* This information could be valuable for healthcare providers in tailoring support and interventions for mothers experiencing mental health concerns.

**Conclusion**:

The graph provides a comparative analysis of maternal mental health status—categorized as stable and concerns—across different gestational ages. The data indicates that maternal mental health concerns tend to fluctuate throughout the gestational period, with notable variability at certain weeks. This suggests that specific stages of pregnancy may be more critical for monitoring and addressing maternal mental health. The presence of error bars highlights the variability within each gestational age group, emphasizing the importance of personalized care and attention to maternal mental health throughout pregnancy.

*  **Age Distribution of Survey Respondents:**

This displot represents the age distribution of survey respondents. The x-axis shows the Age of respondents, ranging from 5 to 45 years, while the y-axis indicates the Count of respondents in each age group. The bars represent the number of respondents within specific age ranges, and the overlaid line graph illustrates the trend of the distribution across different ages. This visualization helps understand the demographic spread and identify age-related trends within the survey data.

**Key observation-**

1. **Axes**:

* X-axis: Represents age, ranging from 20 to 45 years.
* Y-axis: Represents the count of occurrences for each age group.

2. **Displot**:

* The bars represent the frequency of individuals within specific age ranges. The histogram is displayed in light blue with black edges, allowing for clear visibility of the counts.
* The height of each bar indicates the number of individuals in that age range, with peaks indicating more common ages.

1. **Kernel Density Estimate (KDE):**

* The blue line overlaying the histogram represents the KDE, which provides a smoothed estimate of the distribution of ages.
* The KDE helps to visualize the overall trend and shape of the age distribution, highlighting areas of higher density.

1. **Data Interpretation:**

* The histogram shows that the age distribution is relatively uniform, with several peaks around the ages of 25, 30, and 40.
* The KDE line indicates that the most common ages are around 30 and 40, suggesting a concentration of individuals in these age groups.

1. **Trends:**

* The distribution appears to have a slight bimodal shape, with two noticeable peaks, which may suggest that there are two distinct groups within the dataset.
* The counts are relatively stable across the age range, with no significant drop-offs, indicating a diverse age distribution.

1. **Clinical Relevance:**

* Understanding the age distribution can provide insights into the demographics of the population being studied, which can be important for healthcare planning and resource allocation.
* This information may help identify age-related trends in health outcomes or service utilization.**Conclusion**:The graph provides a comprehensive view of the age distribution among survey respondents, highlighting key trends and patterns. The histogram bars show that the majority of respondents fall within the age ranges of 20 to 30 and 35 to 40, indicating these age groups are the most represented in the survey. The superimposed line graph further emphasizes these peaks, suggesting that the survey’s target audience or the most engaged demographic lies within these age brackets. This insight can be valuable for tailoring future surveys, marketing strategies, or services better to meet the needs and preferences of these age groups.

**OVERALL CONCLUSION**

This study provides a comprehensive analysis of pregnancy demographics data collected from *1977 to 2005*, stored in pregnancy demograph datasets. By examining a diverse range of demographic variables, including age, ethnicity, socioeconomic status, and health indicators, we have identified significant trends, patterns, and disparities that can inform healthcare policies and practices.

Our findings underscore the critical role of demographic characteristics in determining pregnancy outcomes. Younger and older age groups were found to have higher risks of complications, highlighting the need for age-specific healthcare strategies. Individuals from lower socioeconomic backgrounds faced greater challenges in accessing quality prenatal care, emphasizing the importance of addressing socioeconomic barriers to improve maternal and neonatal health. Additionally, ethnic disparities in birth outcomes were evident, underscoring the necessity for culturally sensitive healthcare practices and targeted interventions to address the unique needs of diverse populations.

The analysis also revealed the significant impact of maternal health conditions, such as hypertension and diabetes, on pregnancy outcomes. Timely management of these conditions is crucial for improving maternal and neonatal health. Our findings suggest that regular prenatal visits play a vital role in mitigating risks and ensuring positive outcomes. These visits provide opportunities for early detection and management of potential complications, thereby enhancing the overall quality of prenatal care.

By leveraging advanced data analysis techniques, we have uncovered valuable insights that enhance our understanding of pregnancy demographics and their impact on health outcomes. These insights emphasize the need for tailored healthcare strategies to address the specific needs of diverse demographic groups. Healthcare providers and policymakers can use these findings to develop data-driven approaches that enhance the quality of prenatal care and reduce disparities in maternal and neonatal health.

Future research should focus on expanding the dataset to include more variables and a larger sample size. This would provide a more comprehensive understanding of the factors influencing pregnancy outcomes. Additionally, exploring the potential of predictive modeling to forecast pregnancy outcomes could offer valuable tools for healthcare providers to proactively manage risks and improve maternal and neonatal health.

In conclusion, this study contributes to the growing body of knowledge in maternal health and underscores the importance of data analytics in improving healthcare delivery. By identifying key demographic factors and their impact on pregnancy outcomes, we can develop targeted interventions that address the specific needs of diverse populations. This research highlights the potential of data-driven approaches to enhance the quality of prenatal care and reduce disparities in maternal and neonatal health, ultimately contributing to better health outcomes for all.