Smoking and the Developing Effects During Pregnancy and Birth : A North Carolina Study (Circa 04')

Frank Middleton

2024-09-25

Before Starting we run the following :

- This allows for simply data manipulation, creating plots and other visualizations needed or wanted, a color palette to choose from, as well as the data set itself and the system to load it into the R program . We allow for the data from the NC births be read for the following sections.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(RColorBrewer)
library(openintro)
```

```
## Loading required package: airports

## Loading required package: cherryblossom

## Loading required package: usdata
```

```
data("ncbirths")
```

Short Introduction , Description of Data , Variables being explored

In this project we're examining the Birth Data collected from North Carolina from a 2004 random sample conducted. We will go over the birth weights , smoking habits , and how each relates to one another . My thoughts to if their inherently related or if their just individual factors on the same data set, falls to them being connected one way or another. I personally feel like if one was to smoke while being pregnant some sort defect, some inefficiency in development and or timing related .

Variables

- birth weights
- smoking habits
- Pre-Marital Status

Univariate Description of Each Variable Being Considered

Smoking Habits by Birth Weight Category

```
ggplot(ncbirths, aes(x = weight, fill = habit)) +
  geom_bar(position = "dodge", alpha = 0.7) +
  scale_fill_manual(values = c("lightblue1", "darkblue")) +
  labs(title = "Smoking Habits by Birth Weight Category",
       x = "Birth Weight",
       y = "Count",
       fill = "Smoking Status") +
  theme_minimal(base_size = 14)
```

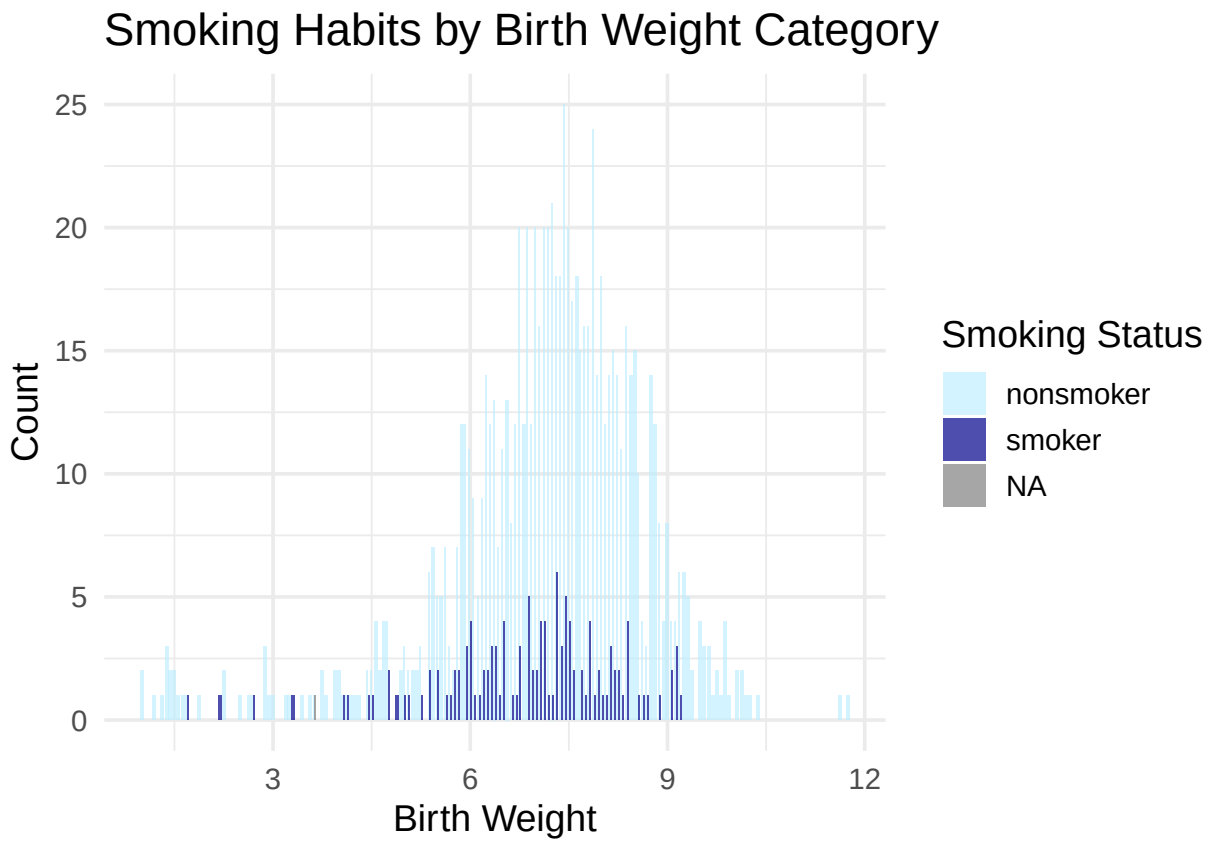

Fig . 1)

The graph (Fig.1) above shows the relative birth weight from everyone serveyed, as well as the count for each weight. Although the Non-smokergroup has a overall majority in total count from the survey there seems to be little differences in mean as the data seems to fall just aroundv 7.3 for the Birth Weight of the babies.

Distribution of Smoking Status

```
smoking_counts <- ncbirths %>% count(habit)

ggplot(ncbirths, aes(x = habit, fill = habit)) +
  geom_bar() +
  scale_fill_manual(values = c("lightblue1", "#72B2D1")) +
  geom_text(stat = "count", aes(label = after_stat(count)), vjust = -0.5, size = 5) +
  labs(title = "Distribution of Smoking Status", x = "Smoking Status", y = "Count") +
  theme_minimal(base_size = 14)
```

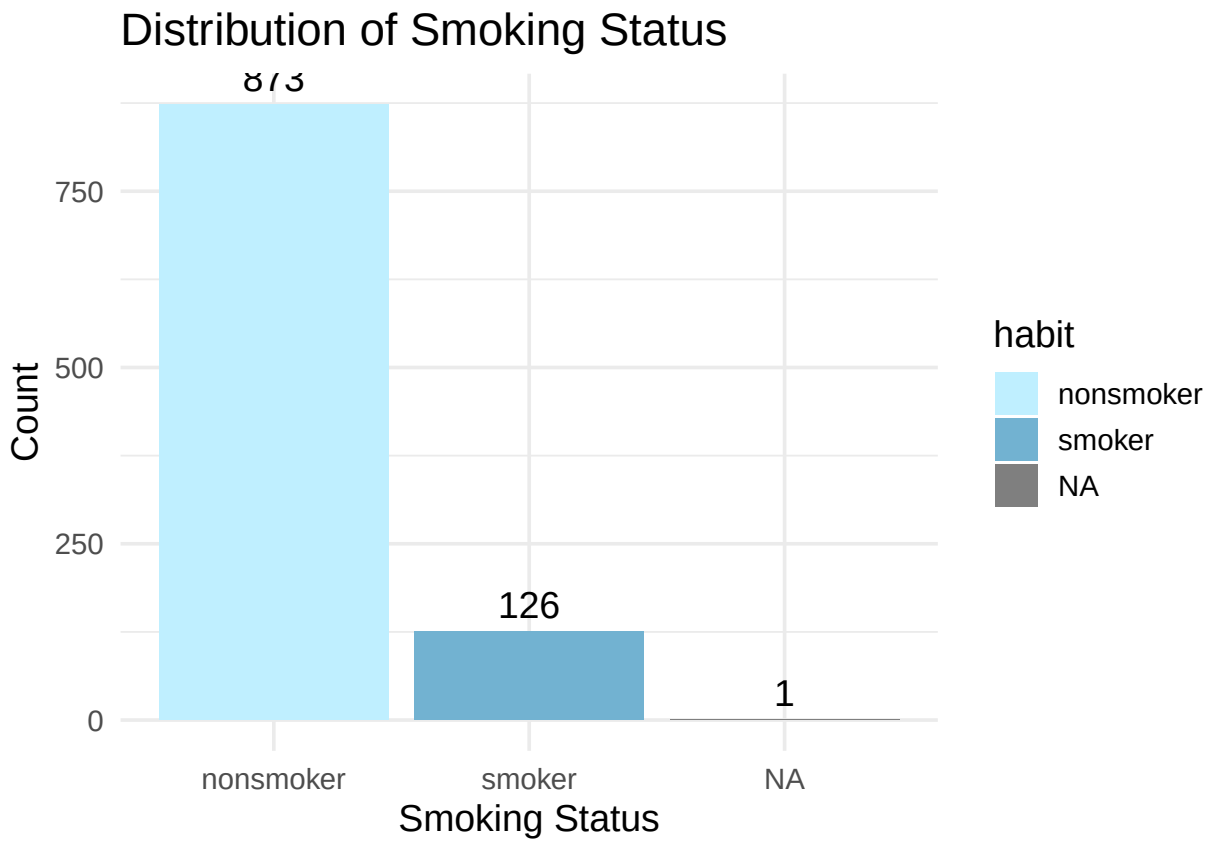

Fig . 2)

The Bar graph above (Fig.2) shows the count of mothers that were either smokers or non-smokers. There seems to be a(n) over majority of people that fall under the non-smoker category where (92%) of the survey can be found the other (8%) can be found above the smokers category.

Distribution of Pre-Maturital Status

```
premie_counts <- ncbirths %>% count(premie)
ggplot(ncbirths, aes(x = premie, fill = premie)) +
  geom_bar() +
  scale_fill_manual(values = c("#6495ED", "#1A237E")) +
  geom_text(stat = "count", aes(label = after_stat(count)), vjust = -0.5, size = 5) +
  labs(title = "Distribution of Prematurital Status", x = "Prematurity Status", y = "Count") +
  theme_minimal(base_size = 14)
```

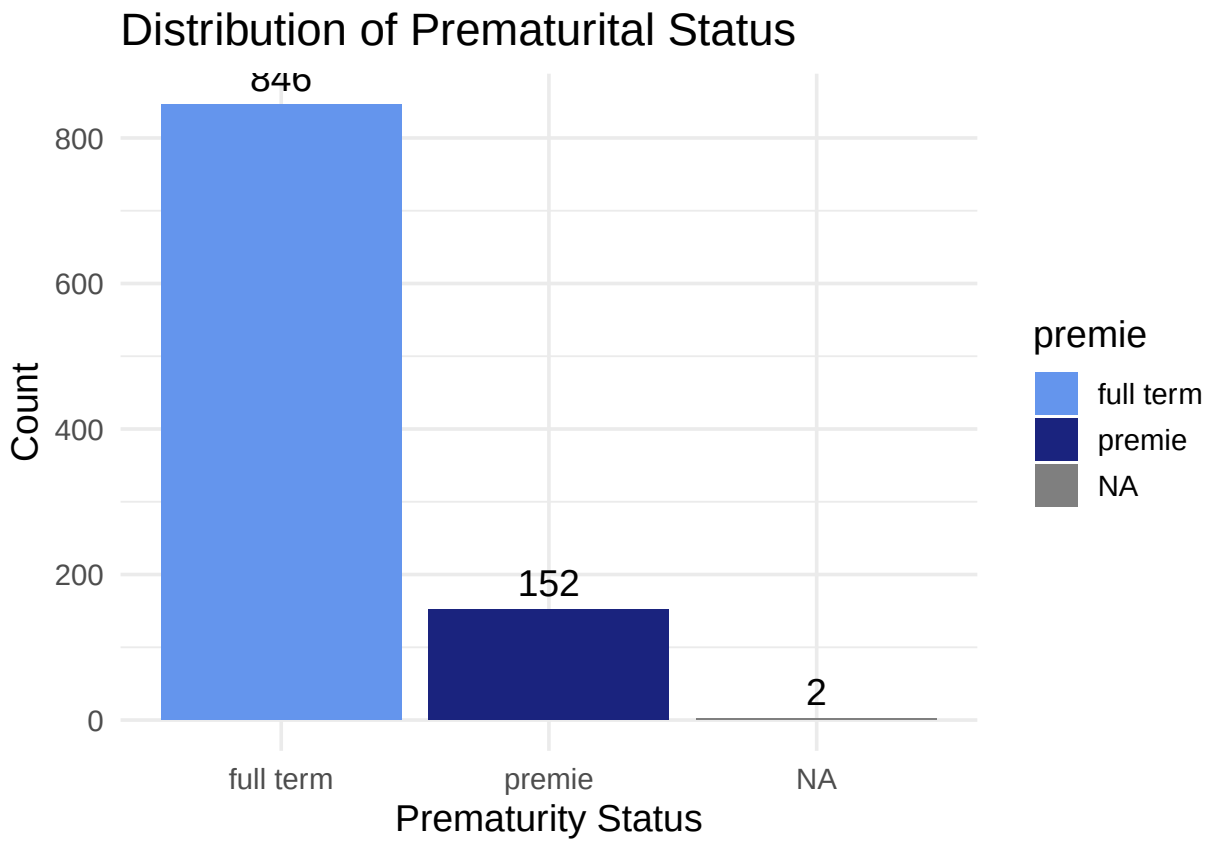

Fig . 3)

The bars above show the simple data collected of the Pre-Maturital Status. This is important as the data is split (93.5%) / (6.5%). As pre mature births are linked to various developmental and health risks, understanding its distribution in relation to the dataset allows us to see further into if there correlated to smoking during pregnacy.

Average Birth Weight by Smoking Habits

```
avg_birth_weight <- ncbirths %>%
  group_by(habit) %>%
  summarise(mean_weight = mean(weight, na.rm = TRUE))

ggplot(avg_birth_weight, aes(x = habit, y = mean_weight, fill = habit)) +
  geom_bar(stat = "identity", position = "dodge") +
  scale_fill_manual(values = c("#004085", "#002")) +
  labs(title = "Average Birth Weight by Smoking Habits",
       x = "Smoking Status",
       y = "Average Birth Weight (grams)") +
  theme_minimal(base_size = 14)
```

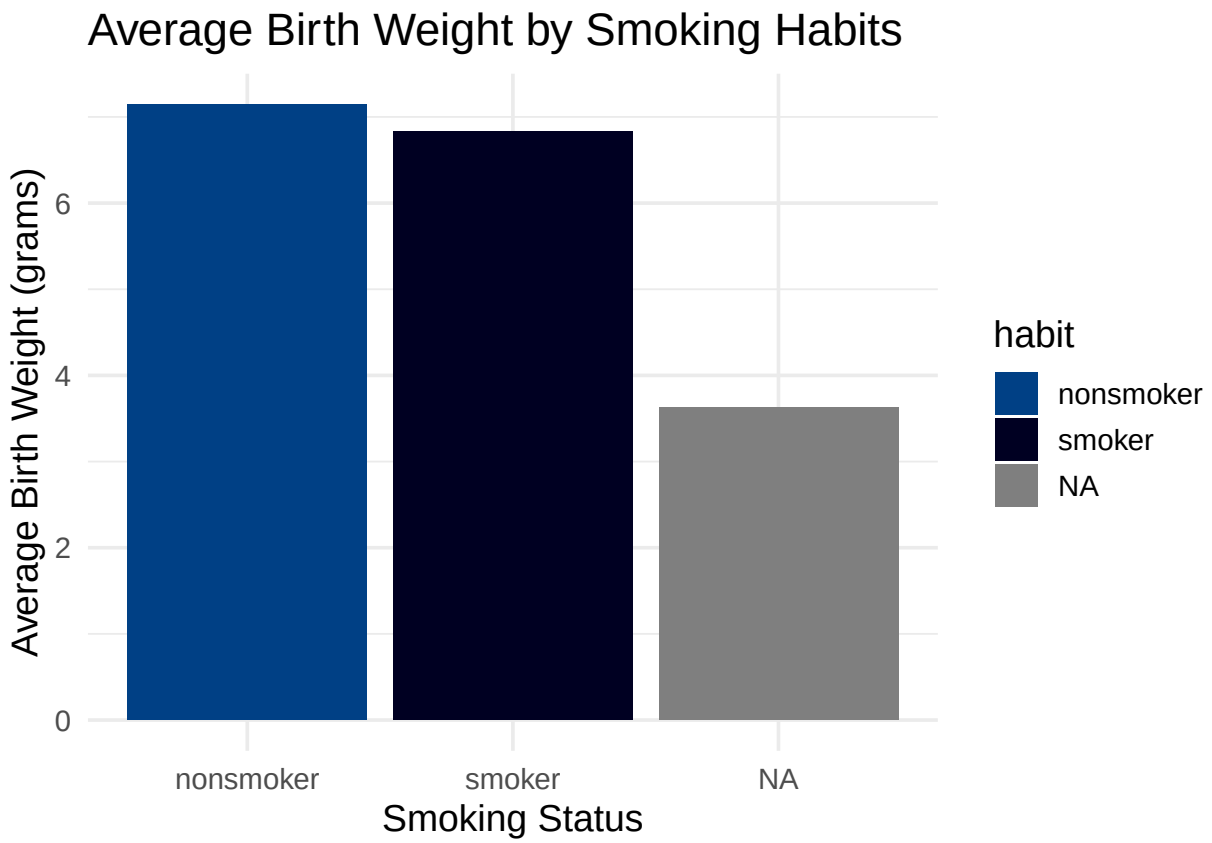

Fig . 4)

This graph compares the average birth weight for infants born to smoking mothers versus non-smoking mothers. The bars are side-by-side to visualize differences in the mean birth weights between the two groups with the mean coming out to be 7.31 . while the current average birth weight is a direct indicator of normal infant health, thelower birth weights are linked to higher infant mortality and morbidity rates overall.

Bivariate Comparison of Two Variables Being Considered

Smoking Habits vs. Prematurity

```
ggplot(ncbirths, aes(x = as.factor(premie), fill = habit)) +
  geom_bar(position = "stack", alpha = 0.7) +
  scale_fill_manual(values = c("#D8BFD8", "purple4")) +
  labs(title = "Smoking Habits vs. Prematurity",
       x = "Prematurity Status (0 = Full-term, 1 = Premature)",
       y = "Count",
       fill = "Smoking Status") +
  theme_minimal(base_size = 14)
```

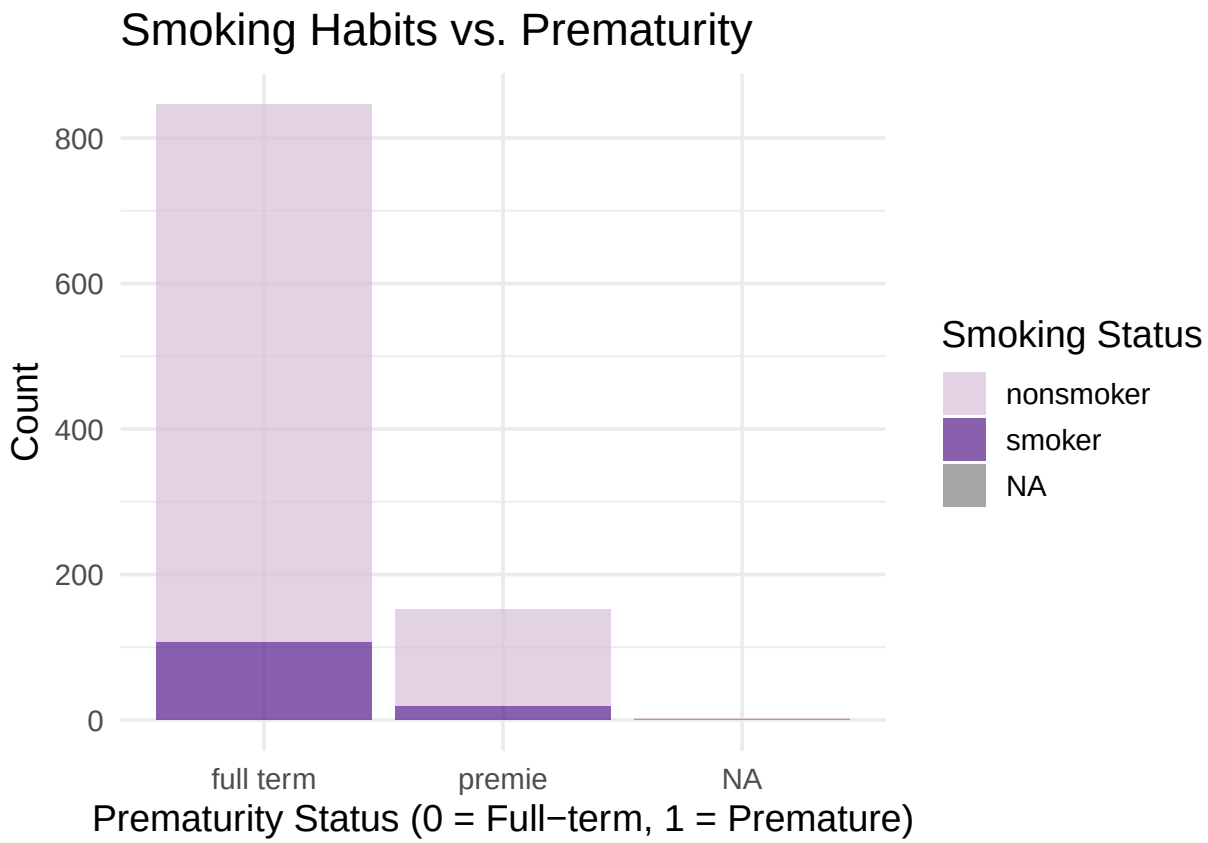

Fig . 5)

This graph (Fig . 5)stacks the bars to compare how smoking habits relate to prematurity. Each bar represents prematurity status (premature vs. full-term). Comparing the two it seems that not only does there seem to be more full term deliveries but roughly the same percentage -(3%)+ . It at this time doesnt seem inherintly corrilated and or presenting a significant hleath concern.

Grouped Summary Statistics for Bivariate Comparison

Smoking Habits vs . Pre-Maturity

```
ncbirths_filtered <- ncbirths %>%
  filter(!is.na(habit) & !is.na(weight) & !is.na(premie)) %>%
  group_by(premie, habit) %>%
  filter(n() > 1)
ggplot(ncbirths_filtered, aes(x = weight, fill = habit)) +
  geom_density(alpha = 0.5, na.rm = TRUE) +
  facet_wrap(~premie) +
  ggtitle("Smoking Habits vs. Prematurity") +
  xlab("Birth Weight (grams)") +
```

```
  scale_fill_manual(values = c("darkblue", "purple4"), name = "Smoking Status") +
  theme_minimal(base_size = 14)
```

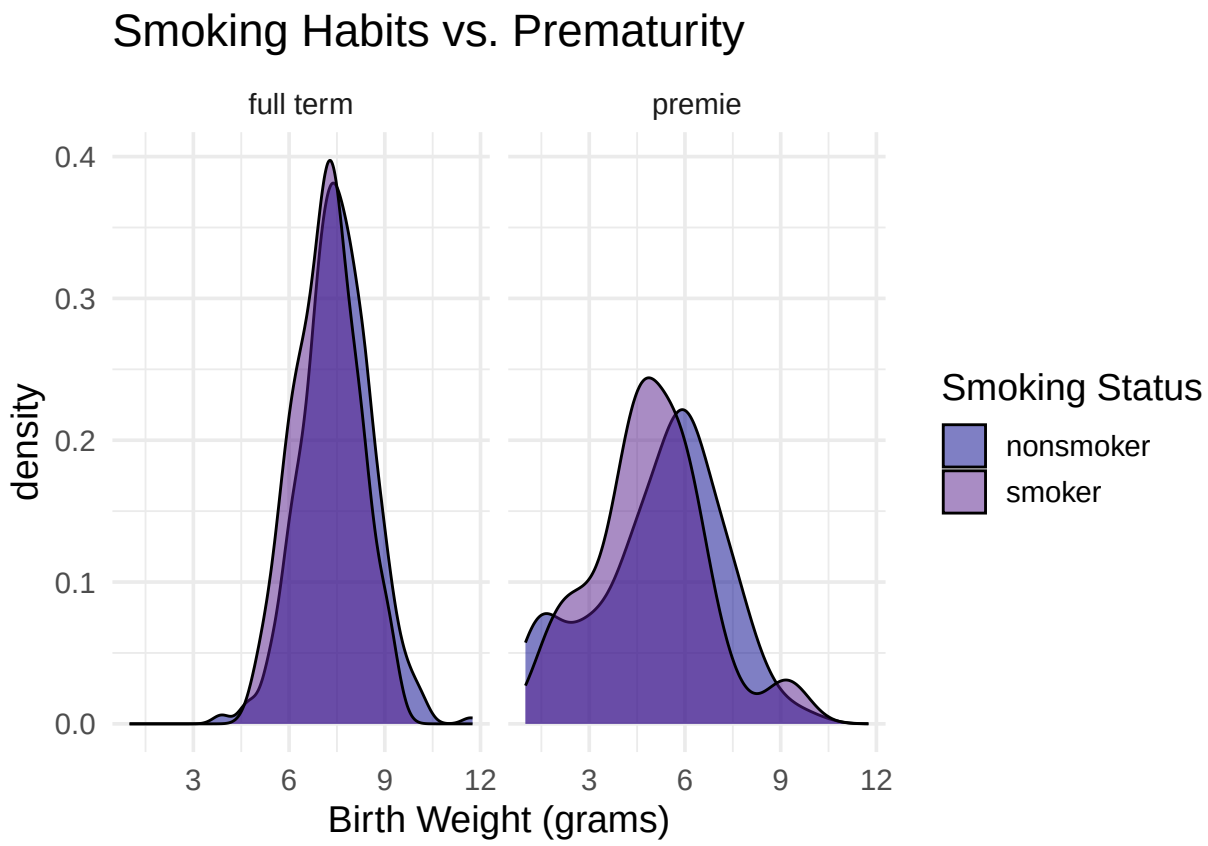

Fig . 6)

This plot shows two graphs of the distribution (density) of birth weights for smoking versus non-smoking mothers, the density plot shows that theres no great significance between the two. From this graph we can see that on average the for the full term hovers around (0.4) , while for the Pre-Mature side averages around (.25). While lower yes there seems to be no great indicators showing any relevance .

Birth weight by Smoking habits

```
ggplot(ncbirths, aes(x = habit, y = weight, fill = habit)) +
  geom_boxplot() +
  scale_fill_manual(values = c("#D8BFD8", "purple4")) +
  labs(title = "Birth Weight by Smoking Habits",
       x = "Smoking Status",
       y = "Birth Weight (grams)") +
  theme_minimal(base_size = 14)
```

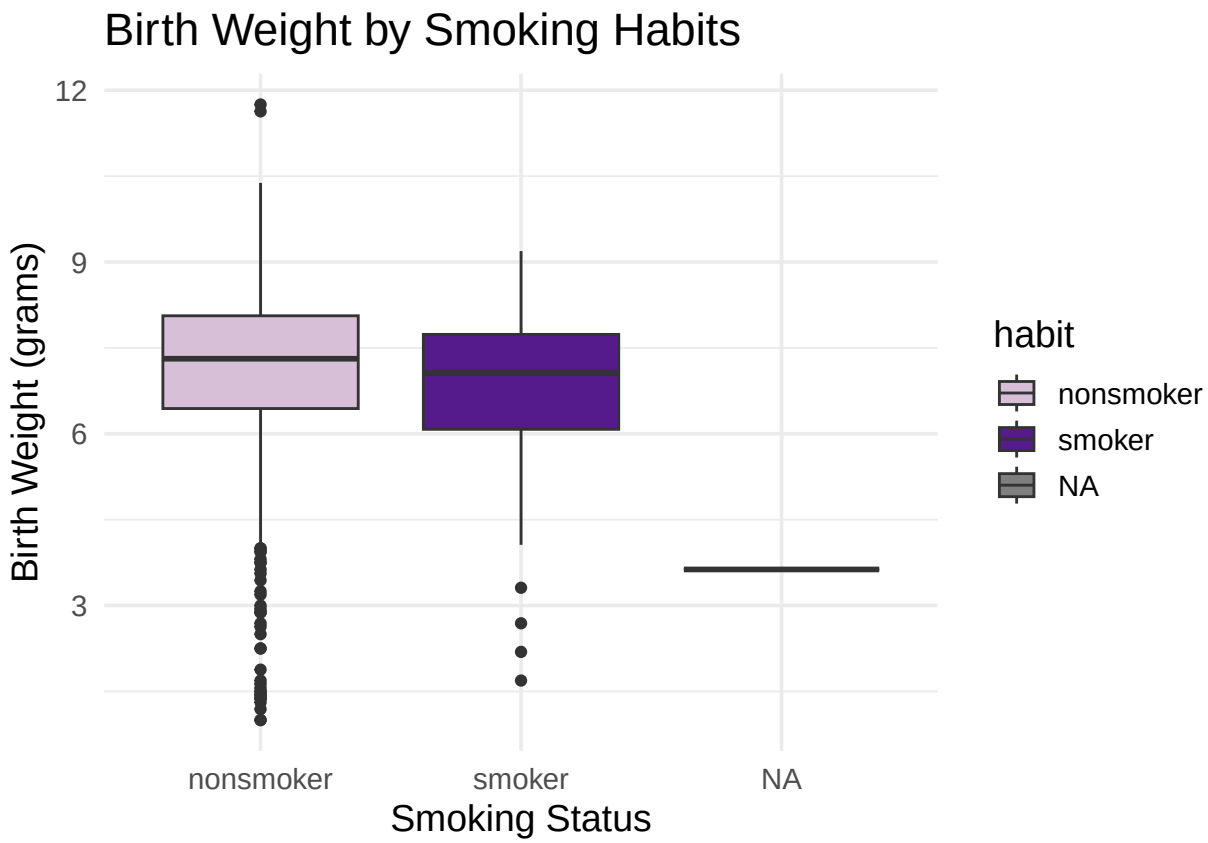

Fig . 7)

```
This box plot compares the birth weights of infants born to smoking versus
non-smoking mothers. While subtle were able to notice that the birth weight
for smokers falls just below the mean avg for the non-smokers group.
There is however a wider range of data for non-smokers   birth weight,
This I believe is due to just the imbalance of data volume for non-smokers has.
```

Conclusion from Univariate and Bivariate Findings

Conclusion

```
In conclusion, the analysis of the 2004 North Carolina birth data reveals that
while there are noticeable differences between smoking and non-smoking mothers,
particularly in birth weight and prematurity status , the overall impact of
tobacco on these variables appears to be subtle. The non-smoker groups
significantly outnumbers the smoker group in the data set, and while the mean
birth weight for infants born to smokers is lower, the difference is not
strikingly significant. The distribution of premature births also shows little
variation between smoking and non-smoking mothers, suggesting that smoking
may not be a direct or sole cause of premature deliveries. These findings imply
that while smoking has a measurable effect on birth weight, the relationship
between smoking, prematurity, and birth outcomes is likely influenced by other
```

factors that require further exploration. Further studies with larger data sets and more comprehensive variables would be necessary to draw stronger conclusions.

While smoking may play a role in reducing birth weight, the data suggests it's not the only factor at play. The complexity of birth outcomes likely involves a range of other influences beyond smoking alone. To gain deeper insights, future research should consider larger datasets and incorporate additional variables to explore how smoking interacts with other factors that could affect pregnancy and birth outcomes. This study provides a snapshot but underscores the need for more nuanced investigation into these relationships.