Project2_B : Color vs Heating/Cooling Dataset

true
true

March 10, 2019

Introduction :-

The excerise is part of taking up various messy data sets and using tidyr and dplyr packages to transform them and infer based on the analysis done. Data preperation/manipulation is the process where data is rearranged, manipulated and prepared for the Analysis to be fed into Model

Problem Statement :-

We have a dataset which has five colors cloths and how they react to colling and heating after every 10 minutes , now this dataset looks pretty simple but it little messyand needs transofmration before we can analyse it and infer something meaning full from it.

Solution :-

The R packages used for the solution are as below.
- **dplyr**
- **tidyr**
- **ggplot2**
- **kableExtra**

1) We load the data using read.csv file and display the raw data using kableExtra library methods in a table using the twitter css styling.
2) Using filter method from dplyr library we seregate the raw_data data frame into 2 new data frames based on the phase condition of ("cooling", "heating").
3) Now using the ggplot, geom_point, geom_line & scale_x_discrete methods from ggplot2 library we plot the color on x-axis and temperature on y-axis and plot showing the effect of cooling and heating on various color (cloth).

```
raw_data <- read.csv("science proj data .csv")

kable(raw_data) %>%
  kable_styling(bootstrap_options = c("striped","hover","condensed","responsive"),full_width   = F,posi
  row_spec(0, background ="gray")
```

color	minute.0	minute.10	minute.20	minute.30	minute.40	minute.50	minute.60	phase
white	78	81	83	88	93	96	98	heating
red	78	82	90	93	98	106	109	heating
pink	78	82	84	90	96	99	102	heating
black	78	88	92	98	108	116	121	heating
green	78	81	85	91	95	102	105	heating
white	98	96	93	80	78	78	78	cooling
red	109	106	95	87	82	80	78	cooling
pink	102	96	90	83	80	78	78	cooling
black	121	108	98	90	84	79	78	cooling
green	105	94	90	82	80	78	78	cooling

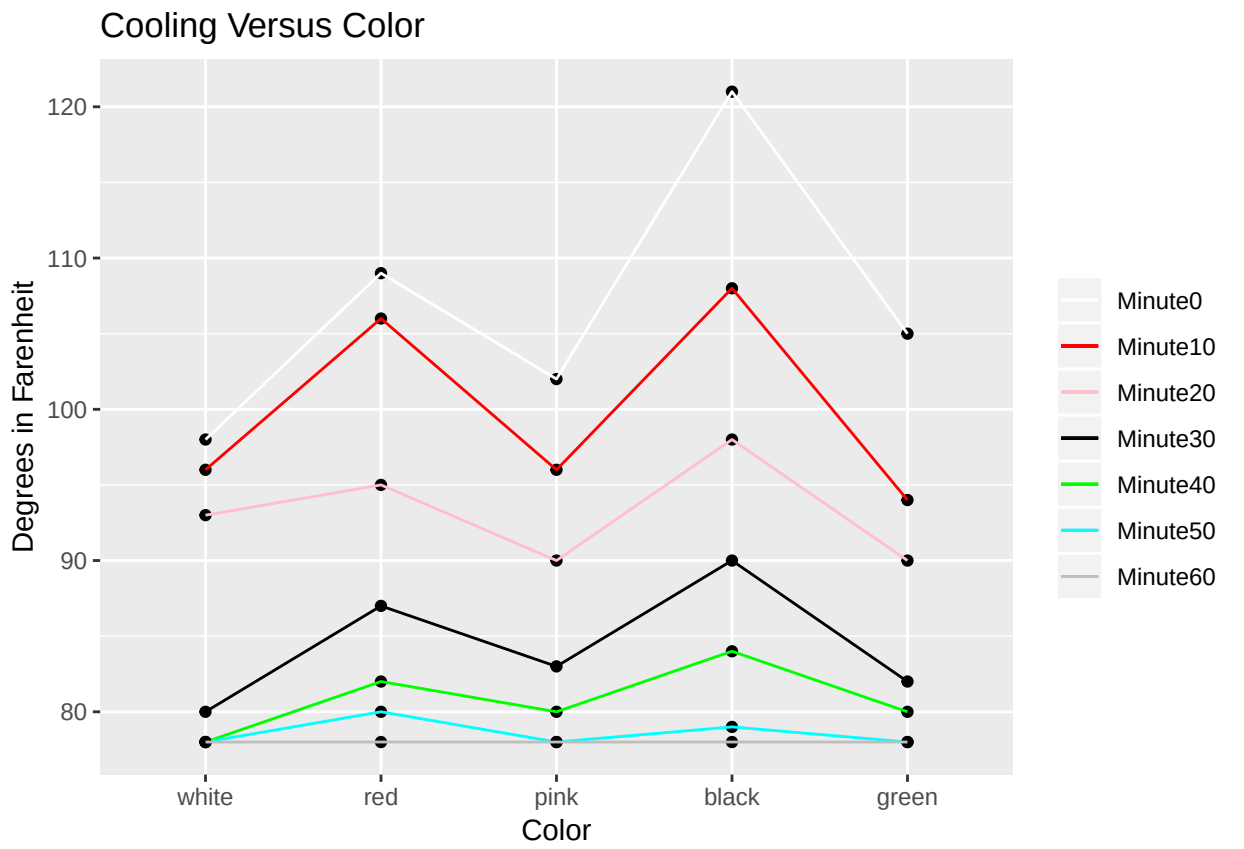
```
df_cooling <- dplyr::filter(raw_data, raw_data$phase == "cooling")

kable(df_cooling) %>%
  kable_styling(bootstrap_options = c("striped","hover","condensed","responsive"),full_width   = F,posi
  row_spec(0, background ="darkgray")
```

color	minute.0	minute.10	minute.20	minute.30	minute.40	minute.50	minute.60	phase
white	98	96	93	80	78	78	78	cooling
red	109	106	95	87	82	80	78	cooling
pink	102	96	90	83	80	78	78	cooling
black	121	108	98	90	84	79	78	cooling
green	105	94	90	82	80	78	78	cooling

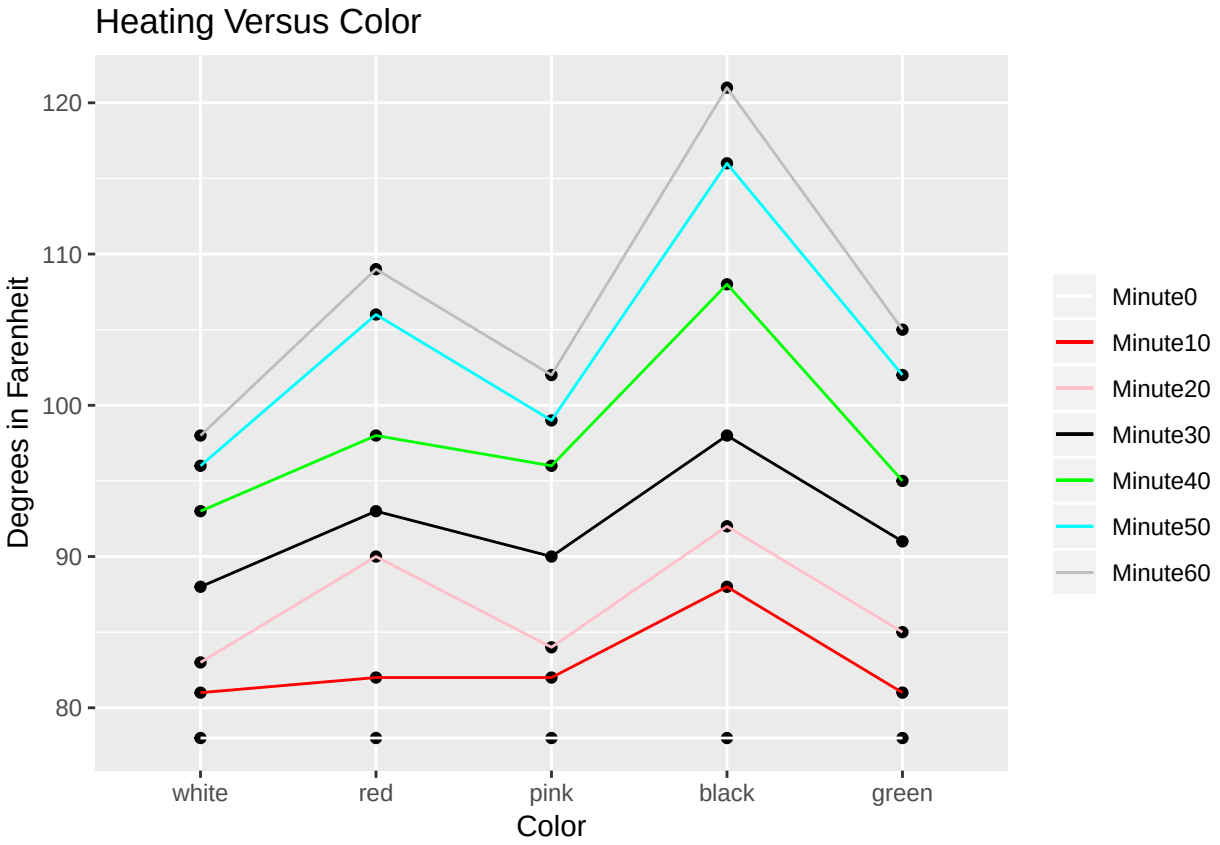
```
df_heating <- dplyr::filter(raw_data, raw_data$phase == 'heating')
kable(df_heating) %>%
  kable_styling(bootstrap_options = c("striped","hover","condensed","responsive"),full_width   = F,posi
  row_spec(0, background ="lightgray")
```

color	minute.0	minute.10	minute.20	minute.30	minute.40	minute.50	minute.60	phase
white	78	81	83	88	93	96	98	heating
red	78	82	90	93	98	106	109	heating
pink	78	82	84	90	96	99	102	heating
black	78	88	92	98	108	116	121	heating
green	78	81	85	91	95	102	105	heating

```
df_cooling %>% ggplot(aes(x=df_cooling$color)) + scale_x_discrete(limits=df_cooling$color) +
  geom_point(aes(y=df_cooling$minute.0 , group=1)) + geom_line(aes(y=df_cooling$minute.0, group=1, col=
  geom_point(aes(y=df_cooling$minute.10, group=1)) + geom_line(aes(y=df_cooling$minute.10, group=1, col=
    geom_point(aes(y=df_cooling$minute.20 , group=1)) + geom_line(aes(y=df_cooling$minute.20, group=1, 
  geom_point(aes(y=df_cooling$minute.30, group=1)) + geom_line(aes(y=df_cooling$minute.30, group=1, col=
  geom_point(aes(y=df_cooling$minute.40, group=1)) + geom_line(aes(y=df_cooling$minute.40, group=1, col=
  geom_point(aes(y=df_cooling$minute.50, group=1)) + geom_line(aes(y=df_cooling$minute.50, group=1, col=
  geom_point(aes(y=df_cooling$minute.60, group=1)) + geom_line(aes(y=df_cooling$minute.60, group=1, col=
  labs(title="Cooling Versus Color", x="Color", y="Degrees in Farenheit", colour="") +
  scale_colour_manual(values = c("white", "red" , "pink" , "black", "green" , "cyan" , "grey"))
```




```
df_heating %>% ggplot(aes(x=df_heating$color)) + scale_x_discrete(limits=df_heating$color) +
  geom_point(aes(y=df_heating$minute.0 , group=1)) + geom_line(aes(y=df_heating$minute.0, group=1, col=
  geom_point(aes(y=df_heating$minute.10, group=1)) + geom_line(aes(y=df_heating$minute.10, group=1, col=
    geom_point(aes(y=df_heating$minute.20 , group=1)) + geom_line(aes(y=df_heating$minute.20, group=1, 
  geom_point(aes(y=df_heating$minute.30, group=1)) + geom_line(aes(y=df_heating$minute.30, group=1, col=
  geom_point(aes(y=df_heating$minute.40, group=1)) + geom_line(aes(y=df_heating$minute.40, group=1, col=
  geom_point(aes(y=df_heating$minute.50, group=1)) + geom_line(aes(y=df_heating$minute.50, group=1, col=
  geom_point(aes(y=df_heating$minute.60, group=1)) + geom_line(aes(y=df_heating$minute.60, group=1, col=
  labs(title="Heating Versus Color", x="Color", y="Degrees in Farenheit", colour="") +
  scale_colour_manual(values = c("white", "red" , "pink" , "black", "green" , "cyan" , "grey"))
```



Summary :-

Thus we can conclude that When a color (colored fabric) absorbs light, it turns the light into thermal energy (heat). The more light a color absorbs, the more thermal energy it produces. Black fabric absorbs all colors of light and is therefore warmer than white fabric which reflects all colors.The same is clearly depicted from the 2 above graphs.