Buoy

```
library(data.table)
library(dplyr)
library(lubridate)
library(ggplot2)
library(zoo)
library(tibble)
library(readr)
```

Getting Buoy Data

```
file_root <- "https://www.ndbc.noaa.gov/view_text_file.php?filename=44013h"
tail <- ".txt.gz&dir=data/historical/stdmet/"

load_buoy_data1 <- function(year) {
  path <- paste0(file_root, year, tail)


  if (year < 2007) {
  header <- scan(path, what = 'character', nlines = 1)
  buoy <- read.table(path, fill = TRUE, header = TRUE, sep = "")
  buoy <- add_column(buoy, mm = NA, .after = "hh")
  buoy <- add_column(buoy, TIDE = NA, .after = "VIS")

  } else {
  header <- scan(path, what = 'character', nlines = 1)
  buoy <- fread(path, header = FALSE, skip = 1, fill = TRUE)

    setnames(buoy, header)
  }

  #return(buoy)
```

```
}

all_data1 <- lapply(1985:2024, load_buoy_data1)

combined_data1 <- rbindlist(all_data1, fill = TRUE)
```

```
load_buoy_data <- function(year) {
  path <- paste0(file_root, year, tail)

  header <- scan(path, what = 'character', nlines = 1)
  num_columns <- length(header)

  if (num_columns == 16) {
    buoy <- read.table(path, fill = TRUE, header = TRUE, sep = "")
    buoy <- add_column(buoy, mm = NA, .after = "hh")
    buoy <- add_column(buoy, TIDE = NA, .after = "VIS")

  } else if (num_columns == 17) {
    buoy <- read.table(path, fill = TRUE, header = TRUE, sep = "")
    buoy <- add_column(buoy, TIDE = NA, .after = "VIS")

  } else {
    buoy <- fread(path, header = FALSE, skip = 1, fill = TRUE)
    setnames(buoy, header)
  }

  return(buoy)
}
all_data <- lapply(1985:2024, load_buoy_data)
combined_data <- rbindlist(all_data, fill = TRUE)
```

```
combined_data1 <- combined_data1 %>%
  mutate(
    YY = as.character(YY),
    `#YY` = as.character(`#YY`),
    YYYY = as.character(YYYY)
  )

# Combine year columns safely using coalesce
combined_data1 <- combined_data1 %>%
  mutate(YYYY = coalesce(YYYY, `#YY`, YY))
combined_data1 <- combined_data1 %>%
```

```
  mutate(BAR = coalesce(as.numeric(BAR), as.numeric(PRES)),  # Convert BAR and PRES to numer
    WD = coalesce(as.numeric(WD), as.numeric(WDIR)))

combined_data1 <- combined_data1 %>%
  select(-TIDE, -TIDE.1, -mm,- WDIR, -PRES,-`#YY`,-YY)

combined_data1$datetime <- ymd_h(paste(combined_data1$YYYY, combined_data1$MM, combined_data

combined_data1 <- combined_data1 %>%
  mutate(across(everything(),
                ~ na_if(as.numeric(as.character(.)), 99) %>%
                na_if(999) %>%
                na_if(9999)))

#summary(combined_data)
#str(combined_data)
#str(combined_data$datetime)
if (!inherits(combined_data1$datetime, "POSIXct")) {
  combined_data1$datetime <- ymd_h(paste(combined_data1$YYYY, combined_data1$MM, combined_da
}
```

```
names(combined_data1)
```

```
 [1] "MM"       "DD"       "hh"       "WD"       "WSPD"     "GST"
 [7] "WVHT"     "DPD"      "APD"      "MWD"      "BAR"      "ATMP"
[13] "WTMP"     "DEWP"     "VIS"      "YYYY"     "mm.1"     "datetime"
```

```
combined_data1 <- combined_data1 %>%
  mutate(Year = year(datetime))

combined_data1 <- combined_data1 %>% select(-YYYY)
```

```
head(combined_data1)
```

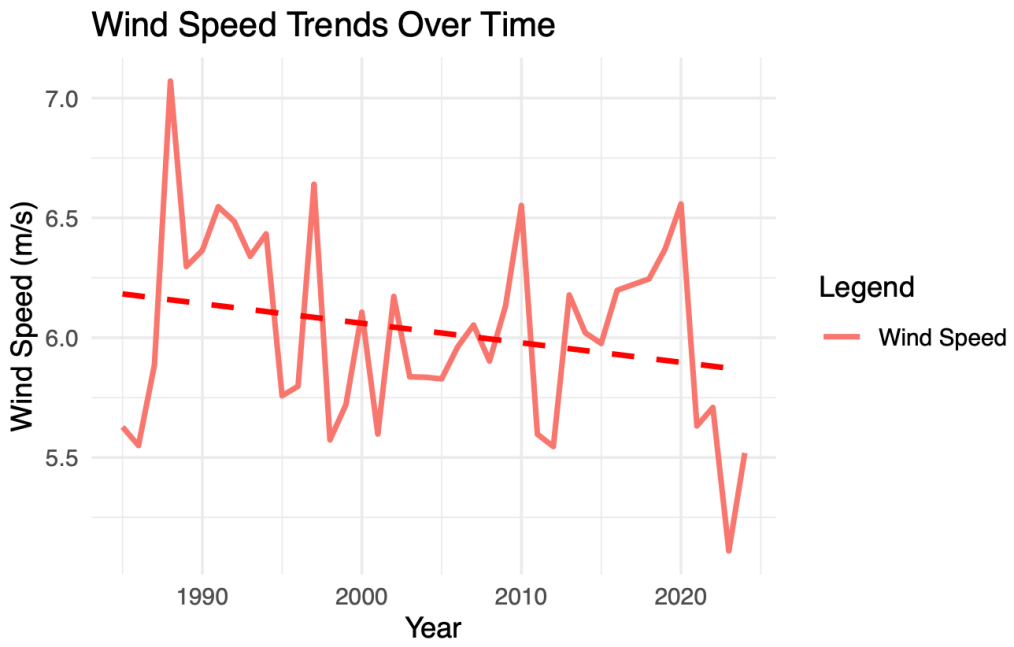
```
      MM    DD    hh    WD  WSPD   GST  WVHT   DPD   APD   MWD    BAR  ATMP
   <num> <num> <num> <num> <num> <num> <num> <num> <num> <num>  <num> <num>
1:     1     1     0    60     4     5    NA    NA    NA    NA 1030.3   4.7
2:     1     1     1    80     4     5    NA    NA    NA    NA 1030.0   5.1
3:     1     1     2   100     4     5    NA    NA    NA    NA 1030.1   5.6
4:     1     1     3   100     4     5    NA    NA    NA    NA 1029.4   5.8
```

```
5:     1     1     4   110     4     5    NA    NA    NA    NA 1028.6   5.8
6:     1     1     5    90     4     5    NA    NA    NA    NA 1027.8   5.3
    WTMP  DEWP   VIS  mm.1             datetime  Year
   <num> <num> <num> <num>               <POSc> <num>
1:   6.7    NA    NA    NA 1985-01-01 00:00:00  1985
2:   6.7    NA    NA    NA 1985-01-01 01:00:00  1985
3:   6.6    NA    NA    NA 1985-01-01 02:00:00  1985
4:   6.7    NA    NA    NA 1985-01-01 03:00:00  1985
5:   6.7    NA    NA    NA 1985-01-01 04:00:00  1985
6:   6.7    NA    NA    NA 1985-01-01 05:00:00  1985
```

Question: Is wind speed increasing or decreasing over time? Can we rely on the trend we see?

```r
yearly_avg_wind_speed <- combined_data1 %>%
  group_by(Year) %>%
  summarise(
    avg_wind_speed = mean(WSPD, na.rm = TRUE)
  )

ggplot(yearly_avg_wind_speed, aes(x = Year, y = avg_wind_speed)) +
  geom_line(aes(color = "Wind Speed"), size = 1) +
  geom_smooth(method = "lm", se = FALSE, color = "red", linetype = "dashed") +
  labs(
    title = "Wind Speed Trends Over Time",
    x = "Year",
    y = "Wind Speed (m/s)",
    color = "Legend"
  ) +
  theme_minimal()
```




```r
# Fit linear regression
model <- lm(avg_wind_speed ~ Year, data = yearly_avg_wind_speed)

# View results
summary(model)
```

```
Call:
lm(formula = avg_wind_speed ~ Year, data = yearly_avg_wind_speed)

Residuals:
     Min       1Q   Median       3Q      Max
-0.76237 -0.30757  0.04197  0.27596  0.91240

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 22.371214  10.738186   2.083    0.044 *
Year        -0.008155   0.005357  -1.522    0.136
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3911 on 38 degrees of freedom
```

```
  (1 observation deleted due to missingness)
Multiple R-squared:  0.05749,   Adjusted R-squared:  0.03268
F-statistic: 2.318 on 1 and 38 DF,  p-value: 0.1362
```

Conclusion

From our model summary we can conclude that there is a slight downward trend in wind speed over the years. Our coefficient shows that wind speed decreases by 0.008 m/s per year. However, this trend is not significant at the 0.05 level as the p value is 0.136. Additionally our R^2 is 0.057 this means only about 5.7% of the variation in average wind speed is explained by year. This is a very weak relationship meaning that year alone is not a strong predictor.