新疆风电

这是新疆 2019 年风电发电及其环境数据

导入数据

```
library(readxl)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts
```

```
library(knitr)
data <- read_excel("./data/新疆风电 2019.xlsx")
# 仅选择前 5 列和前 5 行
data %>%
  select(1:5) %>% # 选择前 5 列
  head(5) %>%
  kable(caption = " 新疆风电 2019 年数据")
```

表 1: 新疆风电 2019 年数据

时间	实际发电功率（mw）	测风塔 30m 风速 (m/s)	测风塔 50m 风速 (m/s)	测风塔 10m 风向 (°)
2019-01-01 00:00:00	0.979591	0	0.000	166.816
2019-01-01 00:15:00	1.150984	0	0.000	166.832
2019-01-01 00:30:00	1.066162	0	0.000	166.859
2019-01-01 00:45:00	0.923717	0	0.000	166.894
2019-01-01 01:00:00	0.813552	0	0.297	166.892

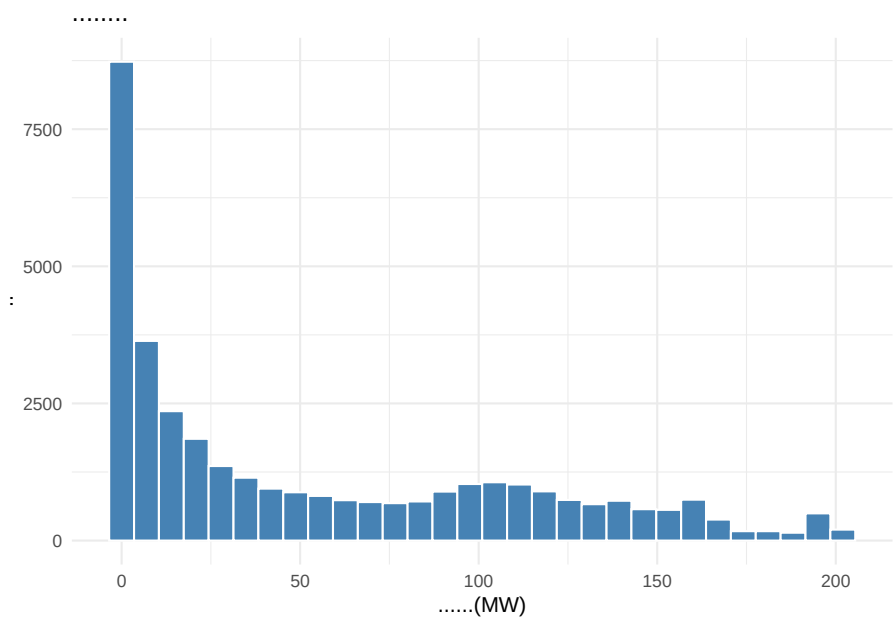
```
# 数据清理
data <- data %>%
  mutate(
    时间 = ymd_hms(时间),          # 转换时间为日期时间格式
    month = month(时间, label = TRUE), # 提取月份
    hour = hour(时间)               # 提取小时
  )
```

可视化

先看看实际发电功率分布

```
ggplot(data, aes(x = `实际发电功率（mw）`)) +
  geom_histogram(bins = 30, fill = "steelblue", color = "white") +
  labs(
    title = " 实际发电功率分布",
    x = " 实际发电功率（MW）",
```

```
    y = " 频数"
  ) +
  theme_minimal()
```



分析

想研究环境对发电的影响，可以从时间、风速、湿度等方面研究

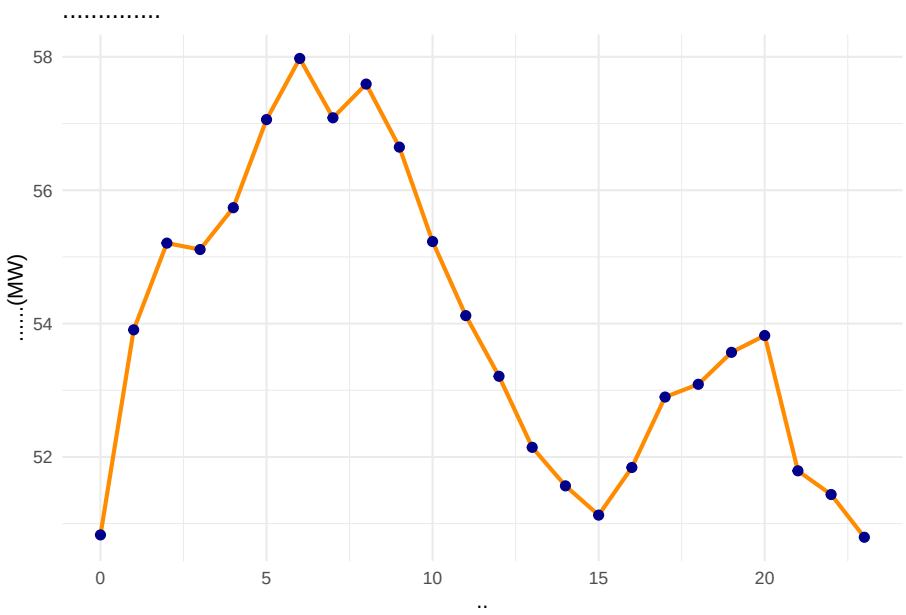
按小时可视化发电功率

```
hourly_data <- data %>%
  group_by(hour) %>%
  summarise(avg_power = mean(`实际发电功率（mw）`, na.rm = TRUE))

ggplot(hourly_data, aes(x = hour, y = avg_power)) +
  geom_line(color = "darkorange", size = 1) +
  geom_point(color = "darkblue", size = 2) +
  labs(
    title = " 一天中不同小时的平均发电功率",
    x = " 小时",
```

```
    y = " 平均发电功率（MW）"
  ) +
  theme_minimal()
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

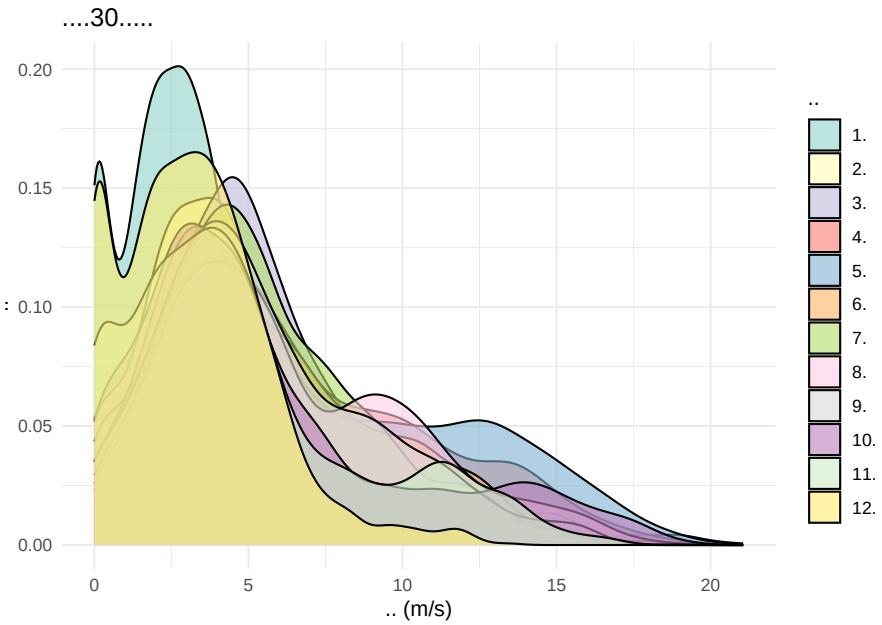


分析

风速的分布

```
ggplot(data, aes(x = `测风塔 30m 风速 (m/s)`, fill = month)) +
  geom_density(alpha = 0.6) +
  scale_fill_brewer(palette = "Set3") +
  labs(
    title = " 不同月份 30 米风速分布",
```

```
    x = " 风速 (m/s)",
    y = " 密度",
    fill = " 月份"
  ) +
  theme_minimal()
```



分析