

新疆 2019 年风电发电

导入数据

```
library(readxl)
library(tidyverse)
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

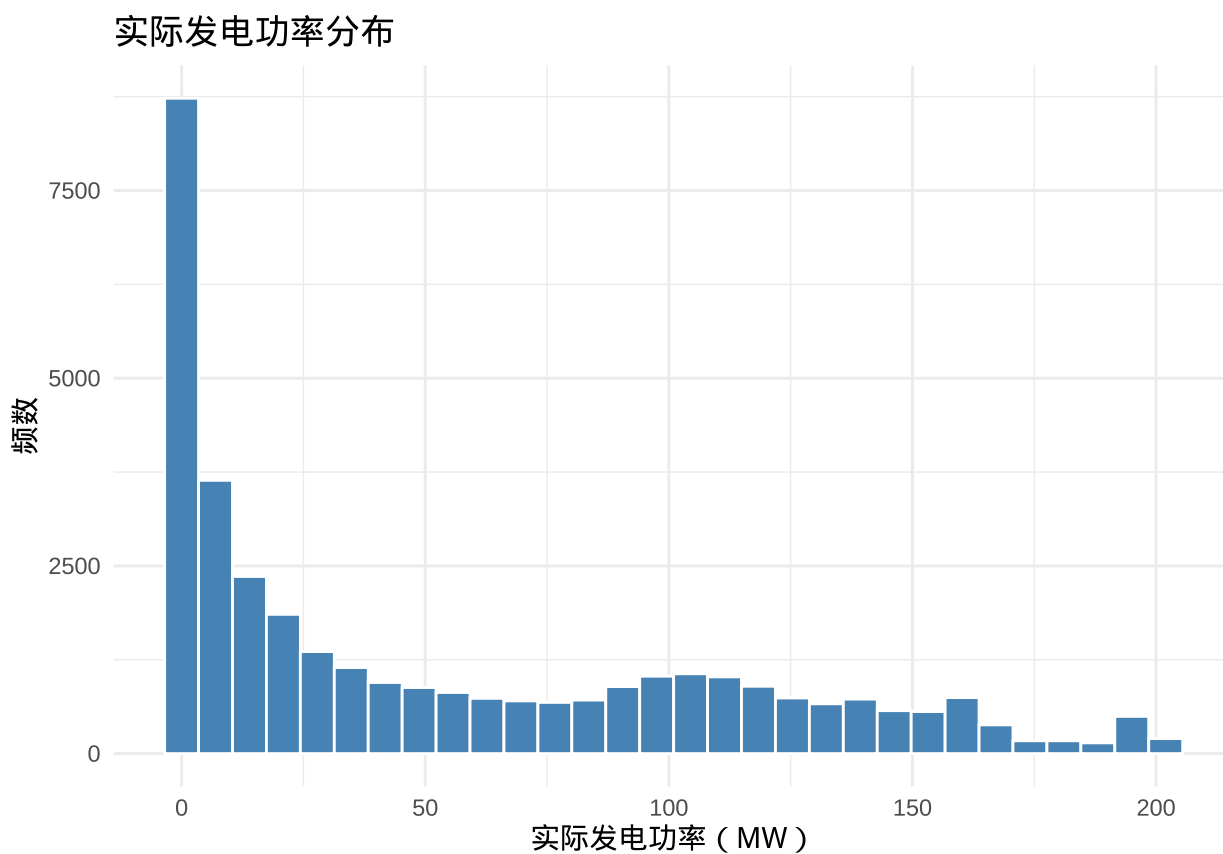
时间	实际发电功率 (mw)	测风塔 30m 风速 (m/s)	测风塔 50m 风速 (m/s)	测风塔 10m 风向 (°)
2019-01-01 00:00:00	0.979591	0	0.000	166.816
2019-01-01 00:15:00	1.150984	0	0.000	166.832
2019-01-01 00:30:00	1.066162	0	0.000	166.859
2019-01-01 00:45:00	0.923717	0	0.000	166.894
2019-01-01 01:00:00	0.813552	0	0.297	166.892

```
# 数据清理, 使用 mutate 函数对 dataframe 类型创建新的列: 时间, month, hour
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

实际发电功率分布

```
# geom_histogram 表示绘制直方图, bins 为柱子的数量, fill 表示填充颜色, color 表示边框颜色
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()
```

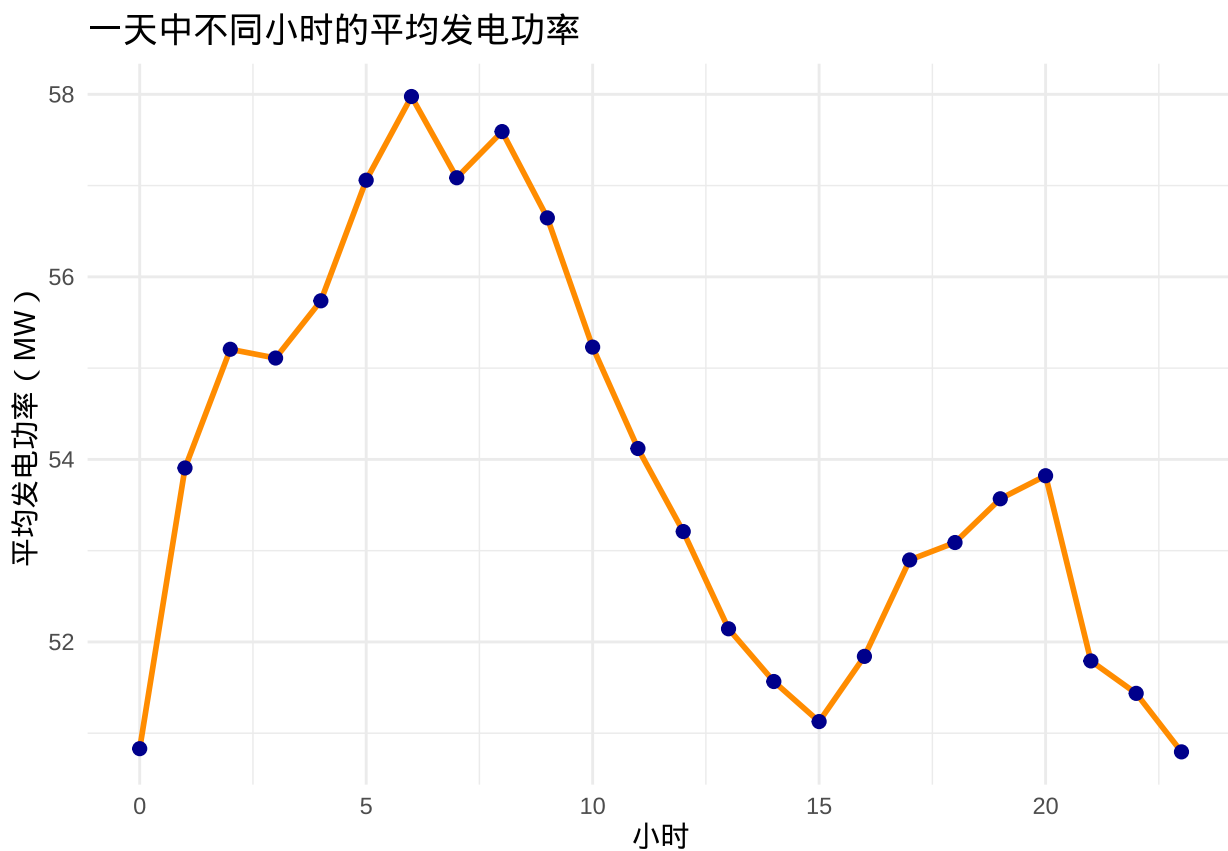


该图展示了实际发电功率的分布情况（单位：MW），横轴为实际发电功率，纵轴为频数。从图中可以看出，发电功率主要集中在低功率范围（0-50 MW），此区间的频数明显最高，特别是功率接近 0 时，频数达到峰值。随着发电功率的增加，频数迅速下降，中高功率范围（100-200 MW）的出现频率较低，分布较为均匀。这种分布可能反映了实际运行中低功率状态的设备使用更为频繁，而高功率状态可能仅在需求峰值或特殊情况下出现。

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

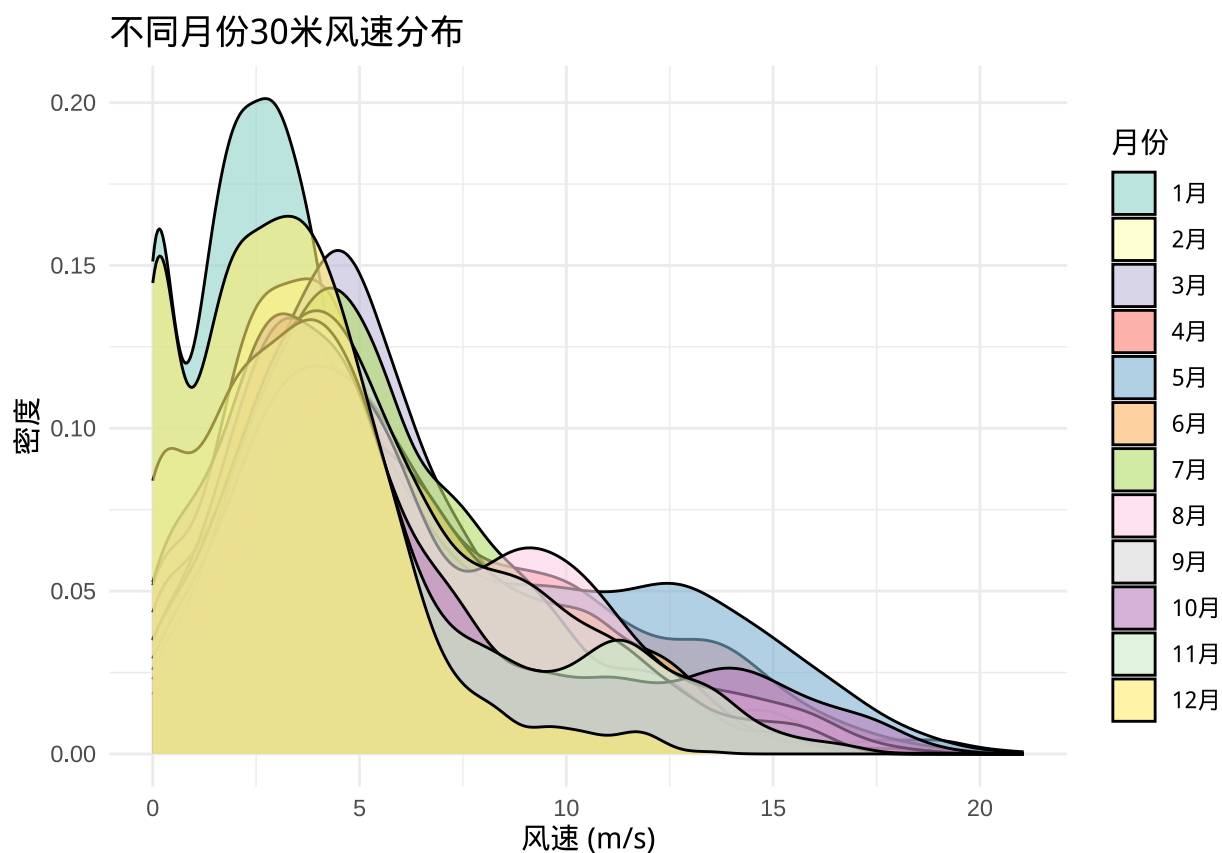
按小时可视化发电功率

```
# geom_line 中设置线段的参数, geom_point 中设置点的参数
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()
```



图中展示了一天 24 小时内发电功率的变化趋势，呈现出典型的日负荷模式。凌晨至早晨（0-6 时）功率逐渐上升，早晨达到高峰（约 58MW）；上午至中午（6-12 时）功率开始下降；下午（12-18 时）继续下降，达到最低点（约 52MW）；晚上时段（18-22 时）功率回升，并在 20 时左右达到次高峰；深夜（22-24 时）功率迅速下降，进入最低点。早晚高峰反映了居民和社会活动的规律，午后低谷和深夜低谷则与人们活动减少和设备关闭有关。### 风速的分布

```
# geom_density 表示绘制密度图, alpha 表示透明度
# scale_fill_brewer 表示用合适的颜色为每个月份的密度图绘制颜色
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()
```



图中展示了不同月份的 30 米高度风速分布特征，风速大多集中在 0~5 m/s 区间，表明低风速发生频率较高，而高风速 (>10 m/s) 则较为罕见。冬季风速分布较窄，集中在低风速区间，反映了天气系统稳定；夏季风速分布较宽，较高风速的概率增大，可能与夏季的对流活动增强有关；春秋季风速分布介于两者之间，表现出过渡特征。随着季节变化，风速分布逐渐扩展，尤其是高风速的发生概率在春夏季略有增加。