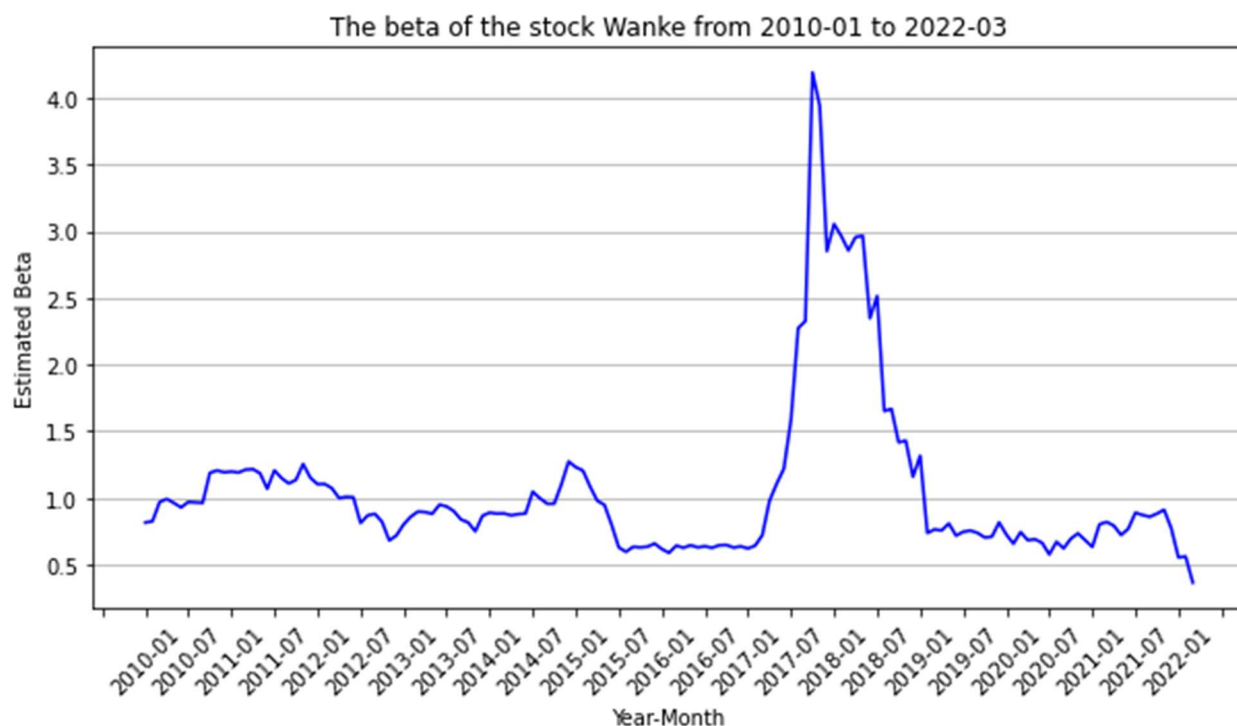


FIN3080 Assignment 2

1. 根据 CSMAR 数据，我做出了如下万科股票的 β 的时间序列图表：



说明：

1) 本图所用数据均来自 CSMAR 数据库，其中万科回报率来自股票市场交易-月个股回报率文件，中证 800 指数回报率来自市场指数-国内指数月行情文件，shibor rate 来自银行体系-shibor 统计表（日）。

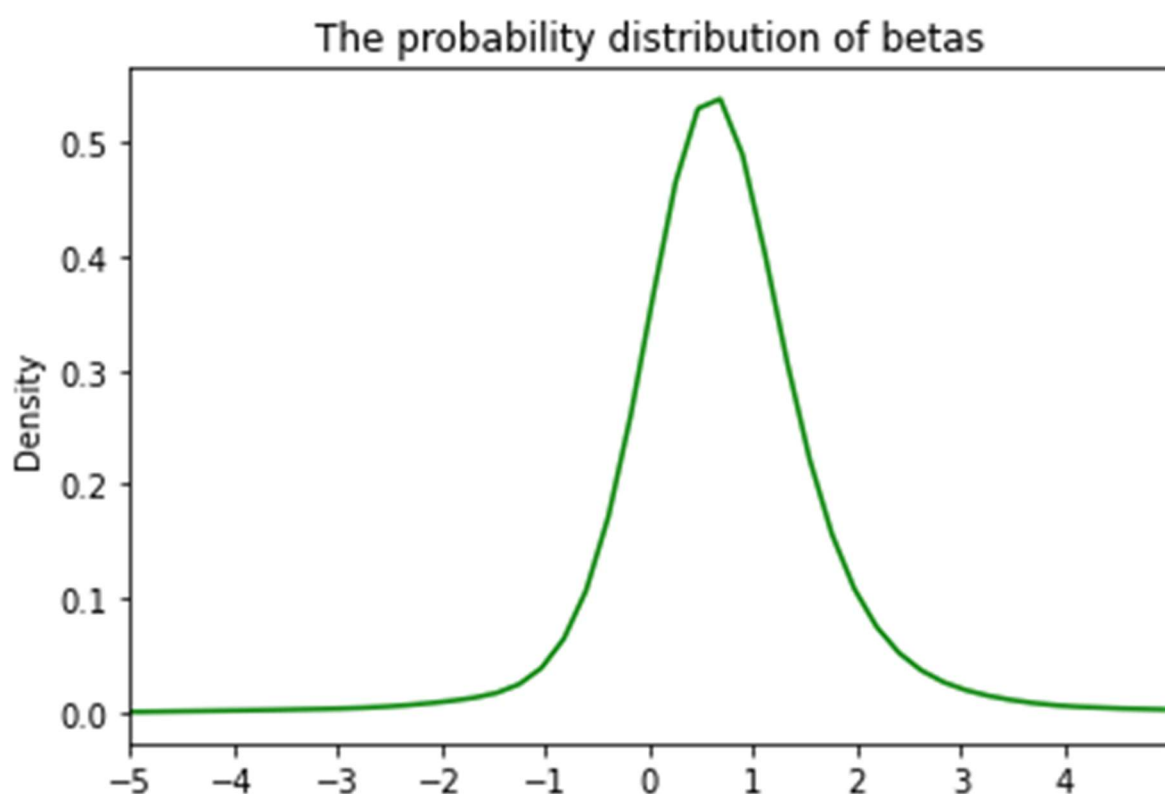
2) 原数据使用字段为 date, Idxrtn, Mretwd, Shibor，经过适当处理使用 statsmodels 的 rollingOLS 模块进行回归，得到 2010-01 至 2022-03 的 beta 数据，后使用 matplotlib 进行图表建立。具体代码文件名为 FIN3080 Homework 2 Q1.ipynb。

3) 对于 2016-01 至 2016-07 缺失值，我采用减少回归项方法处理。若该时期往前推 24 个月数据只有 18 个月数据存在，则采用 18 个月的数据进行回归。数据缺失的原因有可能是万科股权之争导致的该股停牌，导致该段数据缺失。

发现：2017 年至 2018 年底万科股票 β 值较其他时期较高，最高值超过 4。万科作为一

个超大体量的房地产公司，其股票对大盘影响较大，故 β 值应该大于 0，但由于万科在 2015 年开始深陷股权之争直到 2017 年结束，其间由于持股股东抛售或买入大量股票而导致股价波动较大，故其与大盘的协方差较大，造成 β 值在这一时期较大且不稳定。在其他时间阶段与大盘走势相近，由于公司自身原因以及股票市场正常运作造成的股价的小波动体现了 β 值的上下正常浮动。

2. 根据 CSMAR 数据，A 股市场个股的 β 的概率密度分布图以及 β 的百分位值如下：



```
In [10]: perc = [0.1, 0.25, 0.75, 0.9]
df_r.Beta.describe(percentiles = perc)
```

```
Out[10]: count    4078.000000
         mean       0.681220
         std        2.067015
         min       -79.039432
         10%       -0.150774
         25%        0.270248
         50%        0.649266
         75%        1.073994
         90%        1.648505
         max        28.524083
         Name: Beta, dtype: float64
```

说明：

- 1) 个股回报率均来自股票市场交易-月个股回报率文件，大盘数据来自第一题中已经处理完成的数据。
- 2) 对数据进行恰当处理后，使用 sklearn 包中的 LinearRegression 模块进行分组回归，得到个股 2021 年 6 月的预测 β 值，后使用 matplotlib 做出图表，以及 pandas 自带的 describe 函数得到百分位数。具体代码文件名为 FIN3080 Homework 2 Q2.ipynb。
- 3) 最小值与最大值较极端的可能原因是数据量较少，也有公司经营不善导致股价浮动过大的可能。其他分位值均在正常范围之内。
- 4) 所有数据缺失值均执行 drop 操作，对分组回归无影响。