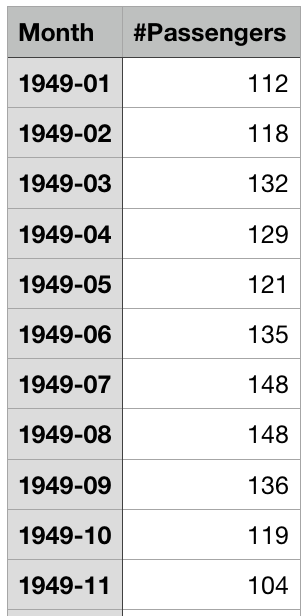
1. 什么是时间序列？

时间序列就是时间间隔不变的情况下收集的时间点上的数值序列。它跟时间无关，不是关于时间的回归，而是纯粹表示随着时间变化自身的变化规律。



1. 几篇不错的博客：
2. <http://www.cnblogs.com/foley/p/5582358.html>
3. <https://www.analyticsvidhya.com/blog/2015/12/complete-tutorial-time-series-modeling/>
4. <http://blog.csdn.net/Earl211/article/details/50957029>
5. <http://www.cnblogs.com/bradleon/p/6832867.html>
6. 关于时间序列的模型有很多，仅从基本模型的角度出发，分为四大类，**AR、MA、ARMA、ARIMA**
7. 所以，面对这些模型，我们该如何是好呢？该如何选择呢？

**我们需要判断时间序列的相关性**，它的相关性有两个主要的系数反应出来，一是**自相关系数（ACF）**，二是**偏自相关系数（PACF）。关于这两个概念的区分，丢个博客：**[**http://blog.csdn.net/wmn7q/article/details/70174300**](http://blog.csdn.net/wmn7q/article/details/70174300)

**第一种情况**，如果ACF函数拖尾（呈几何速率衰减趋向于零），PACF函数截尾（个人理解是在该阶数下PACF特别大，表现出强相关性，而在该阶数之后的PACF都比较小或趋近于零）。**🡪 选择AR模型。**

AR模型：AR（p），其中p表示PACF 在P阶截尾。取AR（P）

**第二种情况**，如果PACF函数拖尾，ACF函数截尾 **🡪选择MA模型**，q为ACF函数相关系数最大的那个点所在的阶数，即MA（q）.

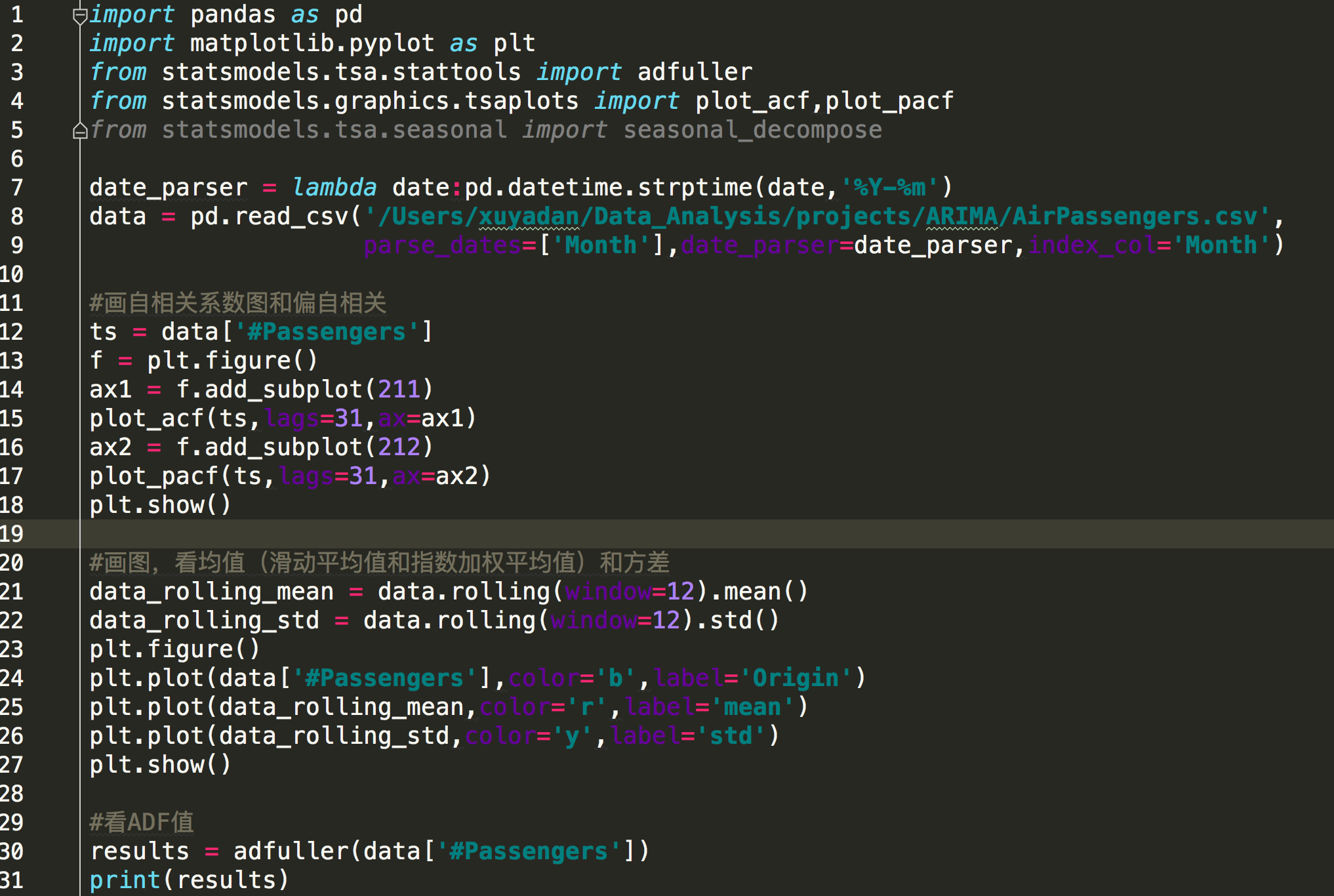
**第三种情况**，如果PACF和ACF全部拖尾，且数列平稳 **🡪选择ARMA模型**；ARMA（p,q）

**第四种情况**，如果PACF和ACF全部拖尾，但是数列不平稳，则需要对不平稳的数列进行差分，使之平稳，然后再继续用ARMA模型，这种情况即为 **🡪ARIMA模型** ARIMA(p,d,q)

1. ARIMA模型的建模步骤涉及相关方法
2. 检测数据是否平稳

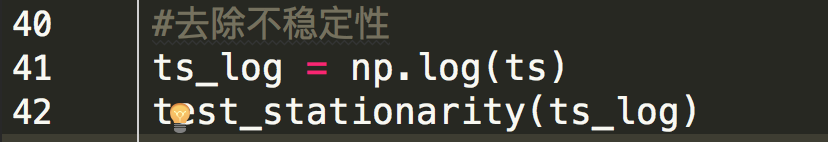
检验数据是否平稳：

有三种方法：画图观察、ADF检验（单位根检验）和看自相关函数和偏自相关函数图



1. 对不平稳的数据进行平稳化处理

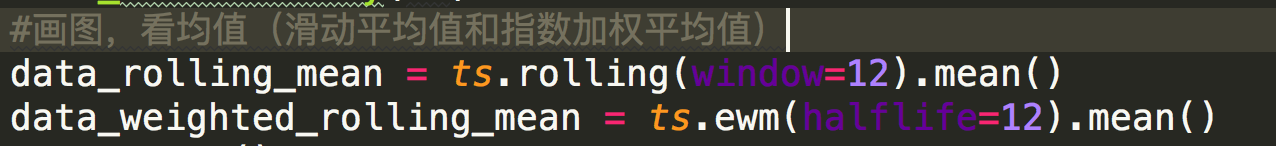
**第一种是对数变换法：**



对数变化相当于加了一个惩罚机制，数据越大惩罚越大，数据越小惩罚越小。缩小了数据的值域。

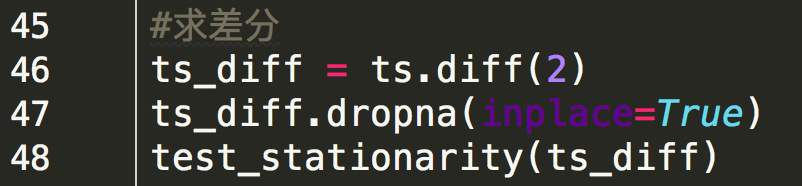
**第二种是平滑法：**

平滑法主要有移动平均法和指数加权平均法。

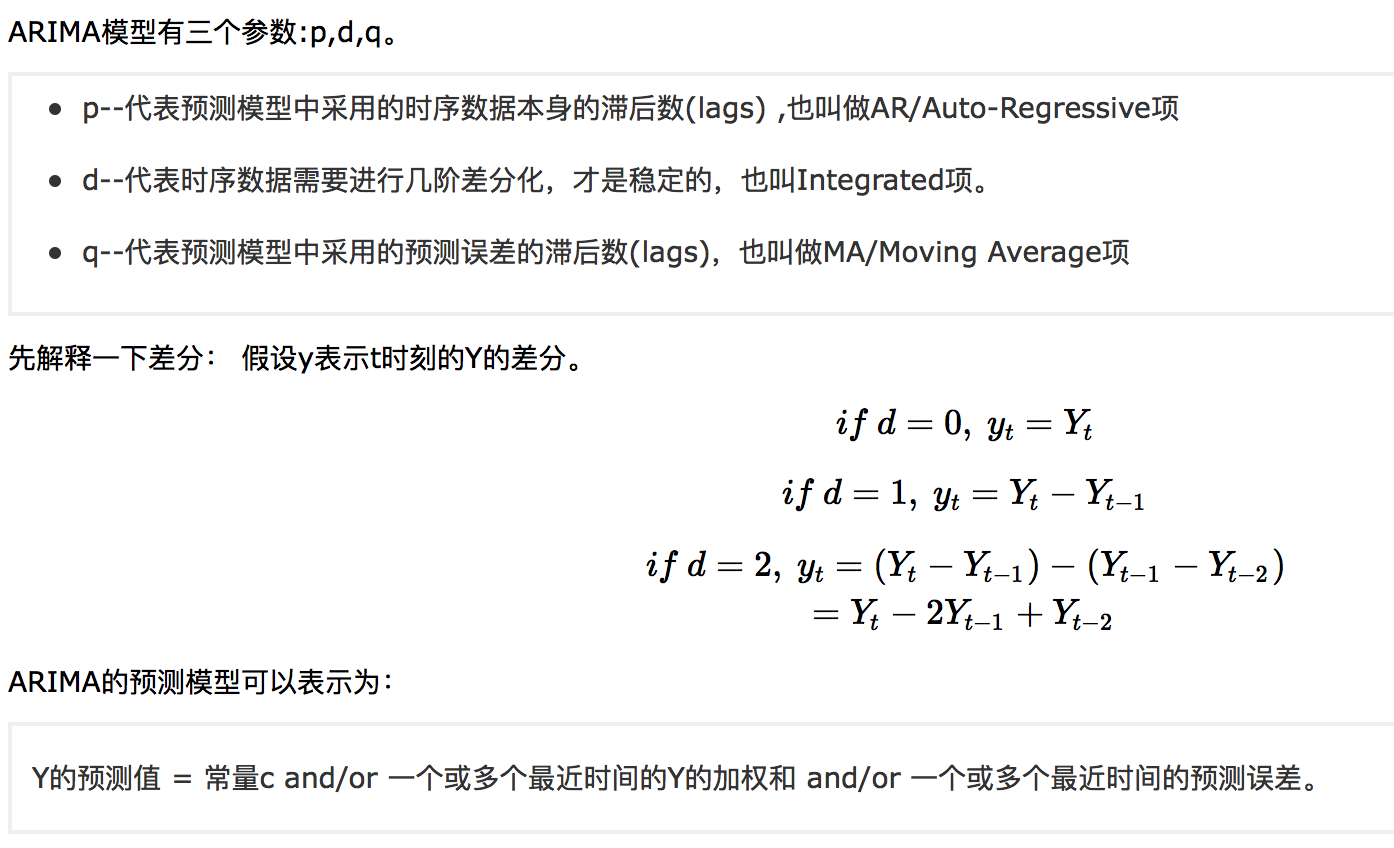


**第三种是差分法：**

也是时间序列数据中用来剔除周期性因素的方法中最常用的方法了。它是对等周期间隔的数据进行线性求减。



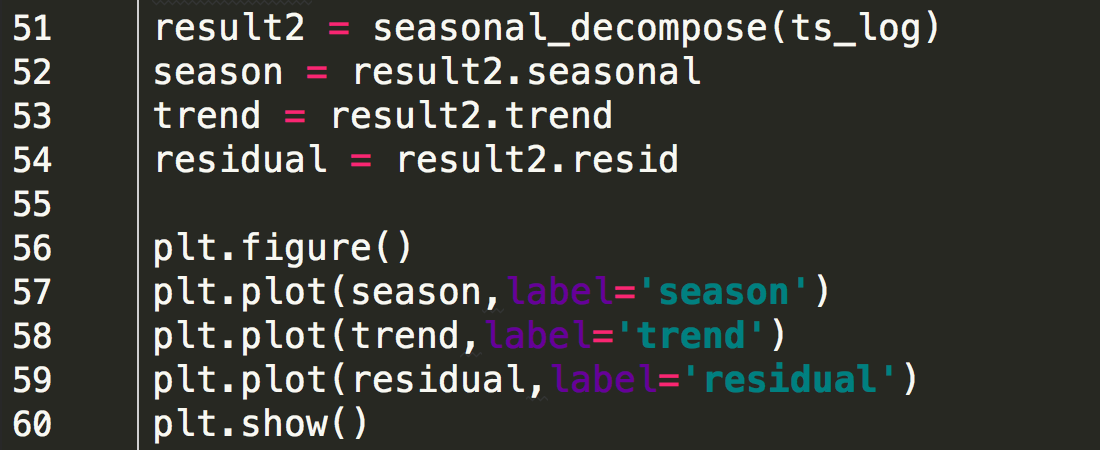
关于差分的解释：



**d表示进行完整的差分的次数。**

**第四种是分解法:**

分解法是将时序数据分解为三个部分：**长期趋势、季节趋势和随机成分（也称为残留成分）**。



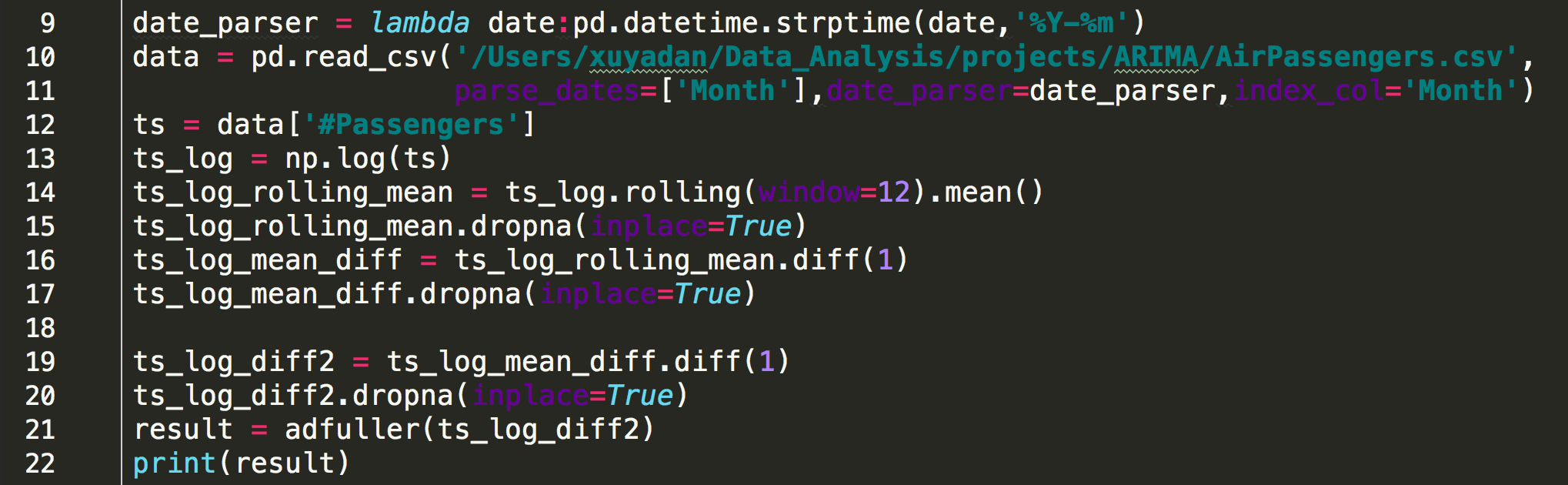
1. 模型拟合：也就是构建模型和参数，fit前面处理过的平稳数据

要进行模型拟合，就需要传入模型所需的参数，例如ARMA（p,q）

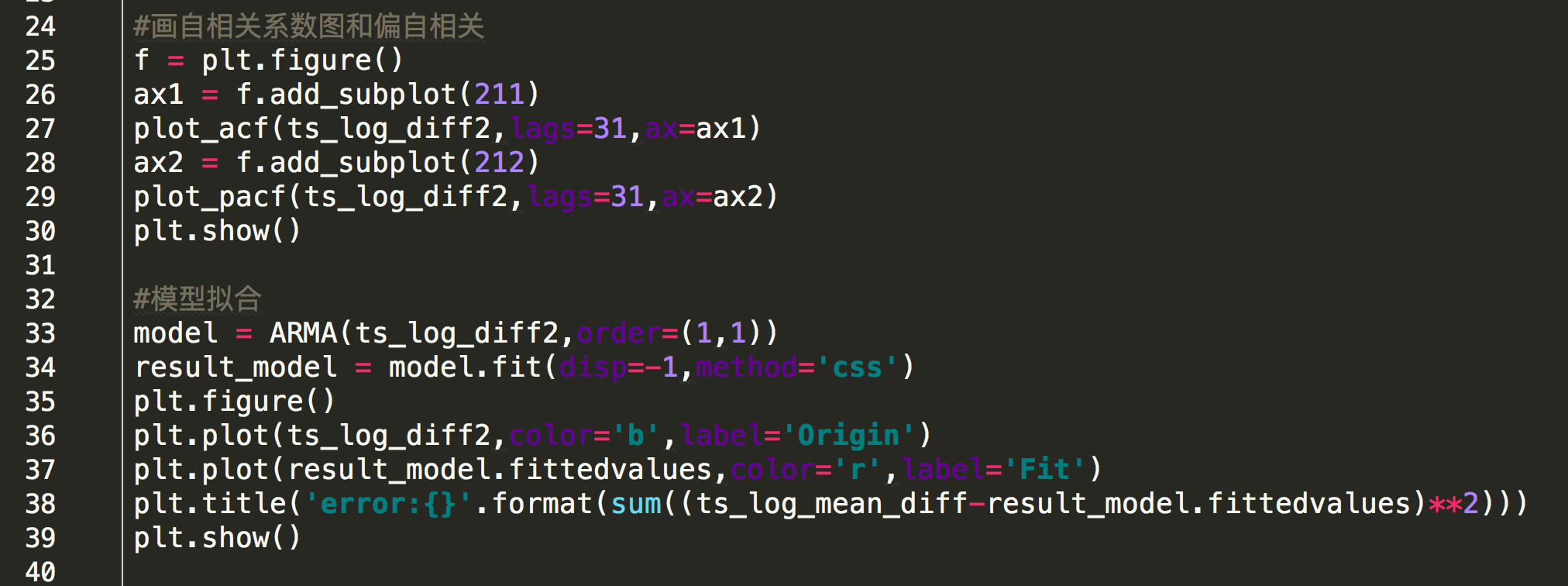
**第一种情况：数据量较小**的情况下，可以通过自相关图和偏相关图来进行模型识别，找出其中的p、q值

**第二种情况：数据量较大**的情况下，就需要通过AIC或者BIC准则来判断模型的好坏，自动从而选出最佳的模型。无论是AIC还是BIC都是最小原则。**AIC考虑模型的参数个数和似然函数，BIC考虑参数数量和样本数量以及似然函数**，BIC相较于AIC有效避免了维数灾难。

1. 模型预测：
2. 完整步骤
3. 数据的处理



1. 找到p,q进行模型拟合



1. 模型的预测

