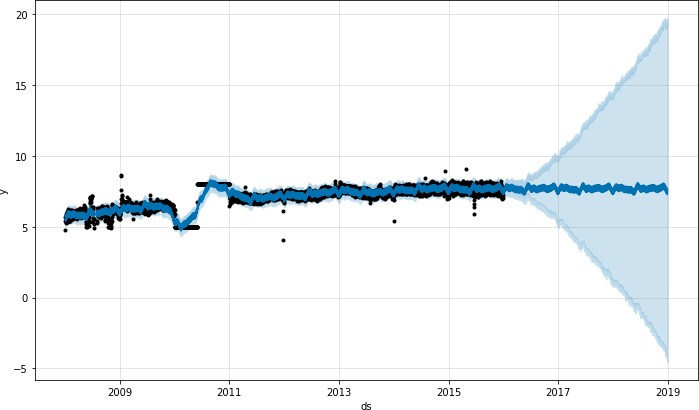
Outliers/ 离群值

# There are two main ways that outliers can affect Prophet forecasts. Here we make a forecast on the logged Wikipedia visits to the R page from before, but with a block of bad data:

# 异常值主要有两种方式影响 Prophet 的预测。在这里，我们对之前记录的维基百科对R页的访问进行了预测，但有一个坏数据块:

1. # Python
2. df = pd.read\_csv('https://raw.githubusercontent.com/facebook/prophet/main/examples/example\_wp\_log\_R\_outliers1.csv') 3 m = Prophet()
3. m.fit(df)
4. future = m.make\_future\_dataframe(periods=1096) 6 forecast = m.predict(future)

7 fig = m.plot(forecast)



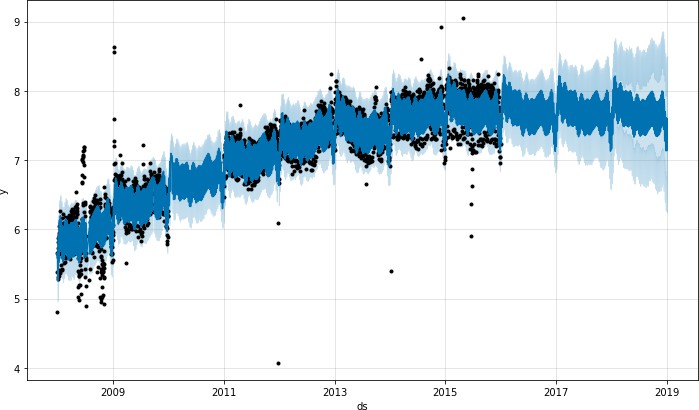
# The trend forecast seems reasonable, but the uncertainty intervals seem way too wide. Prophet is able to handle the outliers in the history, but only by fitting them with trend changes. The uncertainty model then expects future trend changes of similar magnitude. 趋势预测似乎是合理的，但不确定性区间似乎太宽了。Prophet能够处理历史上的异常值，但只能使用趋势变化进行fit。然后，不确定性模型预测未来趋势变化的幅度相似。

The best way to handle outliers is to remove them - Prophet has no problem with missing data. If you set their values to NA in the history but leave the dates in future, then Prophet will give you a prediction for their values. 处理异常值的最好方法是删除它们——Prophet对缺失数据没有任何问题。如果你在历史中设置他们的值为NA，但在未来留下日期，那么先知会给你一个缺失值的预测。

1 # Python

2 df.loc[(df['ds'] > '2010-01-01') & (df['ds'] < '2011-01-01'), 'y'] = None

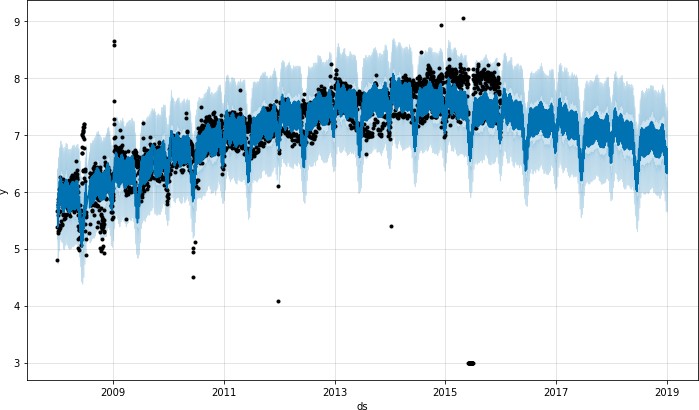
1. model = Prophet().fit(df)
2. fig = model.plot(model.predict(future))



# In the above example the outliers messed up the uncertainty estimation but did not impact the main forecast yhat. This isn’t always the case, as in this example with added outliers:

1. # Python
2. df = pd.read\_csv('https://raw.githubusercontent.com/facebook/prophet/main/examples/example\_wp\_log\_R\_outliers2.csv') 3 m = Prophet()
3. m.fit(df)
4. future = m.make\_future\_dataframe(periods=1096) 6 forecast = m.predict(future)

7 fig = m.plot(forecast)



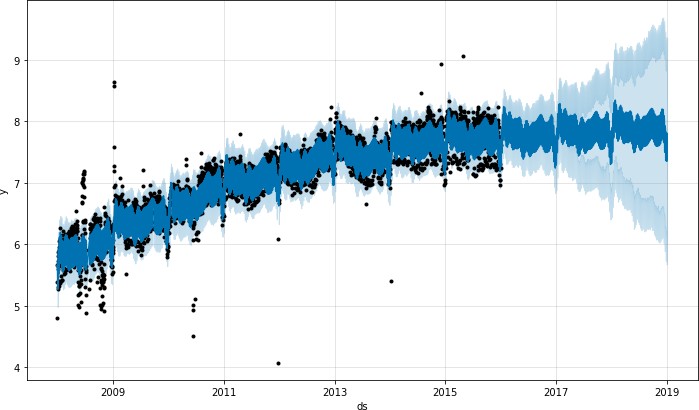
# Here a group of extreme outliers in June 2015 mess up the seasonality estimate, so their effect reverberates into the future forever. Again the right approach is to remove them:

# 在这里，2015年6月的一组极端异常值打乱了季节性估计，因此它们的影响将永远影响到未来。正确的做法是删除它们

1 # Python

2 df.loc[(df['ds'] > '2015-06-01') & (df['ds'] < '2015-06-30'), 'y'] = None

1. m = Prophet().fit(df)
2. fig = m.plot(m.predict(future))



[Edit on GitHub](https://github.com/facebook/prophet/blob/main/docs/_docs/outliers.md)

**Facebook Open Source**

[Open Source Project](https://code.facebook.com/projects/)s [GitHu](https://github.com/facebook/)b [Twitte](https://twitter.com/fbOpenSource)r [Privacy](https://opensource.facebook.com/legal/privacy/) [Terms](https://opensource.facebook.com/legal/terms/) [Contribute to this project on GitHub](https://github.com/facebook/prophet)