基础数据：

1. [cape\_static.csv ]: 船的静态数据，包括imo和对应的mmsi。cape\_match\_static.py
2. [cape\_nav\_201509\_201511.csv] : 船的静态数据 + 停泊事件 + 停泊的港口。
3. [port\_pair\_nav\_AIS.csv] ： 1中的结果提取出港口对，和港口对间的行驶起止时间，提取AIS数据。

计算的中间结果数据：

1. AIS\_cape.py : 在AIS数据中挑出cape船的数据

[AIS\_cape\_20150901.csv-> AIS\_cape\_20150930.csv]

1. AIS\_cape\_all.py：把3中提取出的cape船的AIS数据合并到一个文件中

[AIS\_cape\_ALL.csv] : 把3中提取出的cape船的AIS数据合并到一个文件中

1. NAV\_cape.py : 在停泊事件中挑出cape船的数据

[NAV\_cape\_201509.csv-> NAV\_cape\_201511.csv]

1. cape\_nav.py : 算出每条cape船在3个月内停泊过的港口

[cape\_nav\_201509\_201511.csv]

1. port\_pair.py : 统计cape船在港口间的来往频次 ，2015年9至11月

[port\_pair\_ frequency.csv]

1. port\_pair\_nav.py : cape船港口对的停泊开始结束时间 ，2015年3个月

[port\_pair\_nav.csv]

1. port\_pair\_nav\_AIS.py : 在AIS数据中找停泊开始和结束之间的AIS数据 ：[port\_pair\_nav\_AIS.csv]
2. cleasing\_port\_pair\_nav\_AIS.py : 清洗第7点的结果，把两个相邻数据点距离相差很大的点去掉

[clean\_port\_pair\_nav\_AIS.csv] ：去掉港口对间的数据中，经纬度的距离>速度\*时间的点

修改：

cleasing\_port\_pair\_nav\_AIS.py : 去掉cape的2015.9至11月的AIS数据中，经纬度的距离>速度\*时间的点，并给轨迹分段，两点经度相乘小于0,，就分段，为了把跨越地图两边的横线去掉

[clean\_AIS\_cape\_all.csv]：去掉经纬度的距离>速度\*时间的点

[R\_clean\_AIS\_cape\_all.csv] ：把数据分段，两点经度相乘小于0的点分段，为了在R中画图。画出的是cape船在2015年9月至11月的轨迹图。

1. china\_or\_not.py : 船舶实时分布，判断最近停过的港口是否属于中国 ：[distribution.csv] 2015年11月30日cape船的实时分布，只提取每条船的第一条AIS数据，在[cape\_nav\_201509\_201511.csv]中查找上一次停过的港口，判断是否属于中国。

[China\_real\_distri.csv] 把上个文件拆分，属于中国的船，为了在R中表示，绿色点

[foreign\_real\_distri.csv] 把上个文件拆分，不属于中国的船，为了在R中表示，红色点

1. VLOC\_AIS.py ：提取VLOC船3个月的AIS数据

[VLOC\_imo.csv]：根据给定的VLOC.xls提取出的VLOC船的imo号。

[VLOC\_mmsi.csv] ：根据VLOC的imo在静态数据中提出的对应的mmsi。

[VLOC\_AIS\_201409\_11.csv] ：2014.9至11月的VLOC船的AIS数据

[VLOC\_AIS\_201409\_11\_withGroup2]：每条船一个group。在R中画出来

[VLOC\_AIS\_201509\_11.csv] ：2015.9至11月的VLOC船的AIS数据

[VLOC\_AIS\_201409\_11\_withGroup2] : 用于R画图

1. ave\_speed\_\_day.py : 求每天在航行状态的船只的平均航速。

[speed\_AIS\_cape\_20150901.csv] —> [speed\_AIS\_cape\_20151130.csv] : 把cape船每天的AIS数据，去掉速度小于一节的低速点，去掉加速度大于一节的高速点。剩下的AIS数据

[ave\_speed\_\_day.csv] ：把上面每天的数据；里面的speed字段求一个平均值。每天对应一个平均航速。