

铁路旅客流量分析报告

根据时段、车站、区间等条件分析铁路客流规律

目 录

1

调查背景与目的

2

数据来源

3

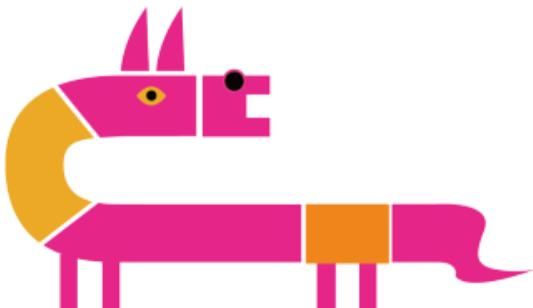
数据采集方式和过程

4

图表数据分析

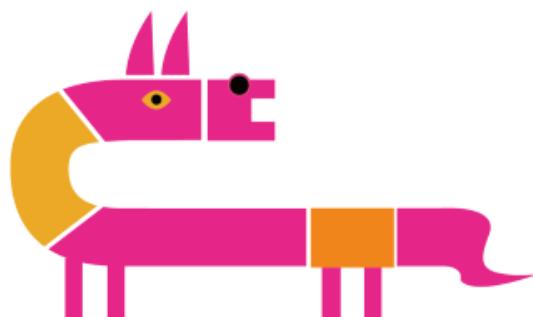
5

总结及建议



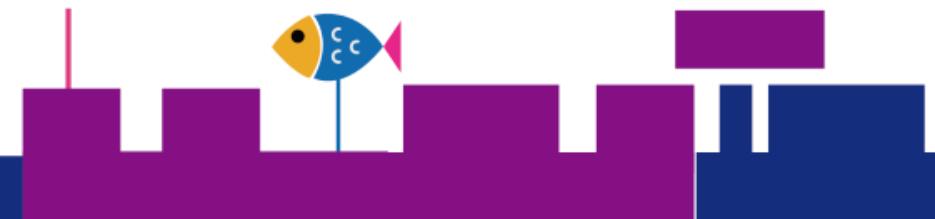
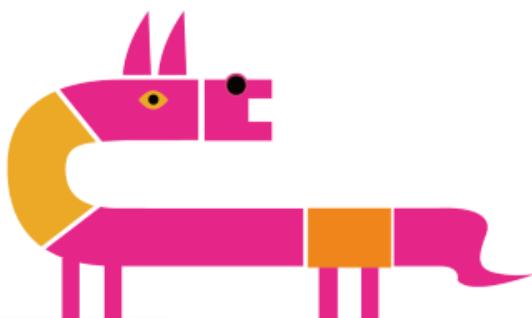
调查背景：

随着发改委发布的《关于改革完善高铁动车组旅客票价政策的通知》，高铁动车票价将由铁路总公司（下称铁总）根据市场情况自行定价的政策出台。铁路部门为了保持市场的竞争力，实现利润的最大化，需要了解日常铁路客运流量、淡旺季变动指数、冷热门线路的具体情况，而其中对客流的充分了解和分析是准确把握市场的首要条件，因此铁路客流预测的研究也成为铁路客运服务需要重点研究的方向。



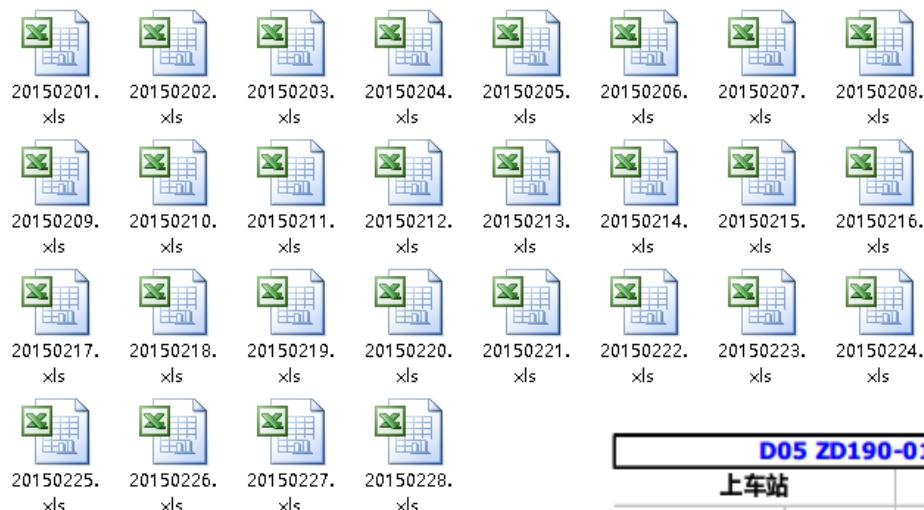
调查目的：

铁路客流量受多种因素的影响，比如：“春运”期间铁路客流量骤增，导致铁路运力无法满足客户乘车需求，同时也给铁路客运组织带来巨大压力。在非节假日期间，一些冷门线路区间上座率不足，造成铁路车辆资源的浪费。因此铁路客流量分析，可以为制定合理的价格、改善客运站组织方式、优化铁路车辆资源配置、提高客运设备的服务能力提供帮助，对提高铁路客运运输效率具有重要的意义。



数据来源：

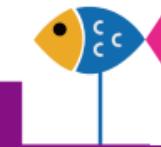
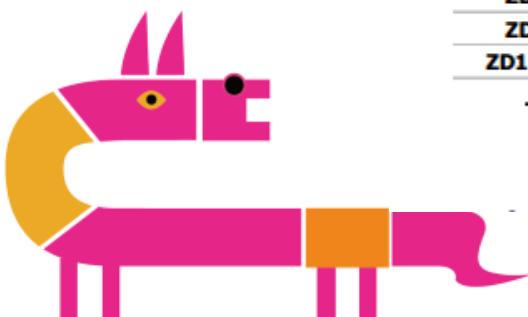
第四届“泰迪杯”全国数据挖掘挑战赛赛题B，包含2015年1月1日至2016年3月20日某铁路局的所有铁路相关信息。



D05 ZD190-01—ZD111-01 开行 1 天 日均定员：1172 客座率：47.3 %					
上车站		ZD190-01	ZD062	ZD250	ZD326
下车站	开到点	06:43	07:30	下车人数合计	
ZD190-01	06:43		17		17
ZD062	07:28		13	28	41
ZD250	07:57		64	18	75
ZD326	08:35		123	62	327
ZD111-01	09:27		217	108	430
上车人数合计		217	108	402	430
1157					

ZD190-01 站上车总人数。

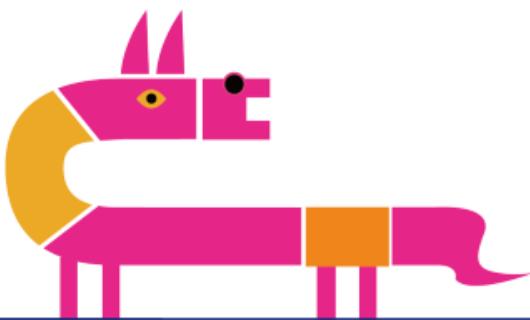
ZD190-01 站上车后的旅客，在沿途各站下车的人数统计



大数据描述：

本次所提供第三方数据主要有：450天每个车次始发日期、日均定员、所经站点、客座率，以及在不同时间点的站点处上下车人数以及总客流
量（备注：excel和html表格形式，相关格式内容可在“数据来源”查看）

数据来自于真实的铁路公司。保障了本次数据分析的真实性、可靠性以及完整性。大幅度减少误差程度以及对数据分析的片面性。

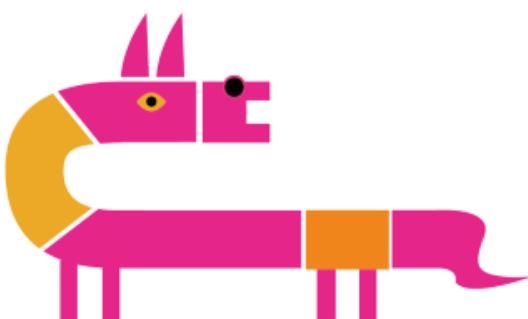


数据采集方式和过程：

1、数据库分析与设计

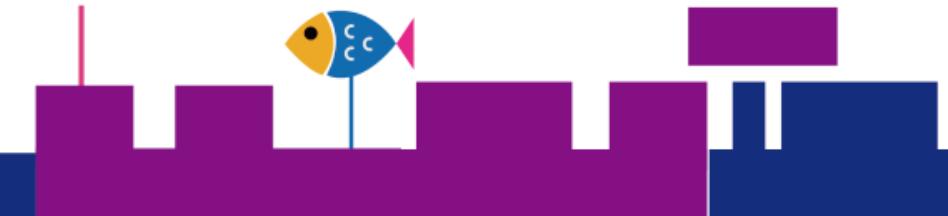
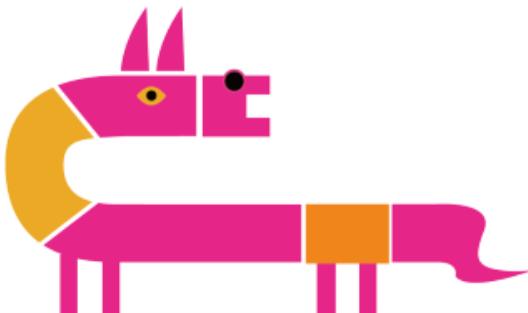
我们采用数据库的方式对采集数据进行存储管理，可便于提取重要信息、分析数据，并且可用于铁路部门对车站客流数据的管理，包括各类数据信息的添加、编辑、查询和删除等。

我们使用了SQL Server 2008开发数据库，可直接对文件进行集中管理，开发时还可以根据需要向其中添加或新建数据库、表、表单和报表等。在各个表中输入数据记录，并在数据库设计器中根据表的共同字段“地区”、“日期”、“站点”、“车次”建立表之间一对一的关系。

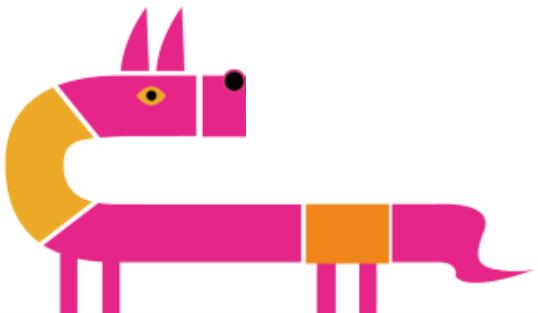
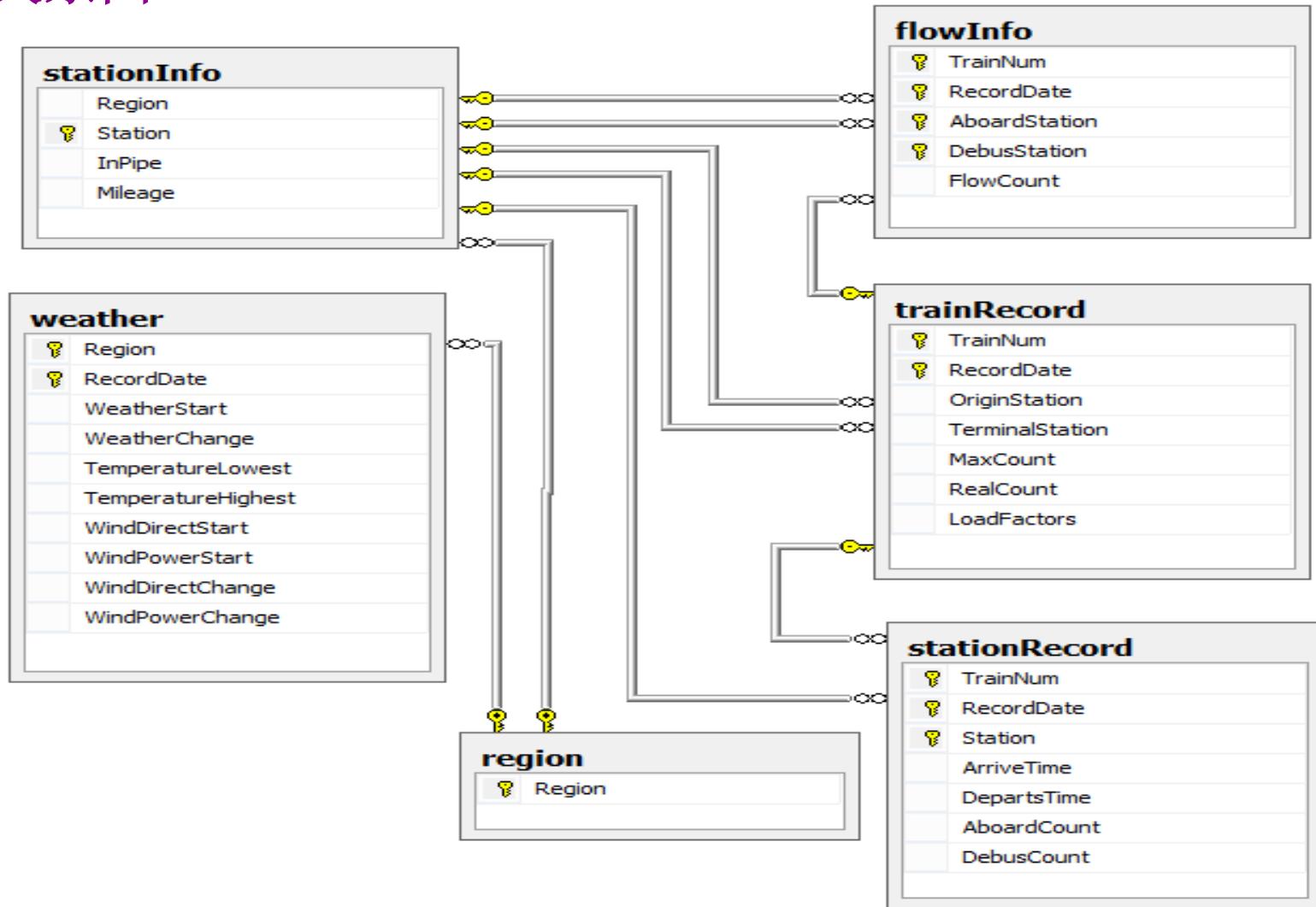


2、功能模块设计

- ① 地区管理模块：地区名。
- ② 站点信息模块：地区、站点。
- ③ 天气状况记录模块：地区、日期、起始天气、转换后天气、最低气温、最高气温、起始风力风向、转换后风力风向。
- ④ 车次记录模块：车次、日期、出发站、终点站、日均定员、实际乘车人数、客座率。
- ⑤ 站点记录模块：车次、日期、当前站点、车次抵达时间、发车时间、上车人数、下车人数。
- ⑥ 客流信息模块：车次、日期、上车站、下车站、乘客数。



3、数据库关系图

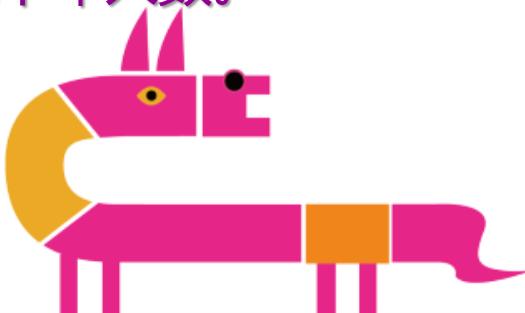


4、数据提取与导入

数据提取采用C#编写提取器进行提取。

首先进行数据的预处理，对于原始的旅客列车梯形密度表，将列车车次、开车日期、出发站点、出发时间、终点站、终到时间、客座率提取出来放到trainRecord.csv文件中，将途径车站、到达时间、出发时间对应车次信息提取出来放到stationRecord.csv文件中，最后将站点间的客流信息提取出来对应车次信息和起始站点放到flowInfo.csv中。

然后将整理出来的数据上传到服务器中，将trainRecord.csv、 stationRecord.csv、 flowInfo.csv中的数据上传到服务器中对应的trainRecord、 stationRecord、 flowInfo表中，由数据库的触发器通过flowInfo的客流人数计算出每个列车的载客人数和在各站点的上下车人数。



列车运行系统数据导入器

文件 地区、车站信息导入 天气信息导入 列车运行信息导入

列车运行系统数据导入器

文件 地区、车站信息导入 天气信息导入 列车运行信息导入

地区信息导入 车站信息导入

列车运行系统数据导入器

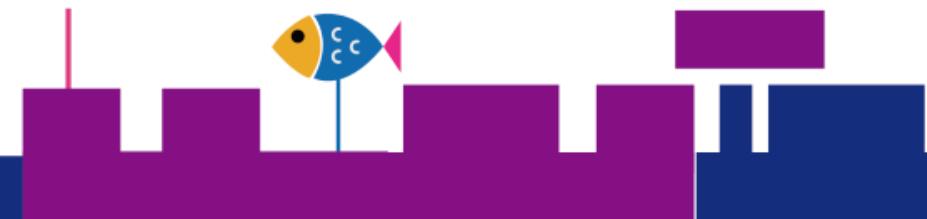
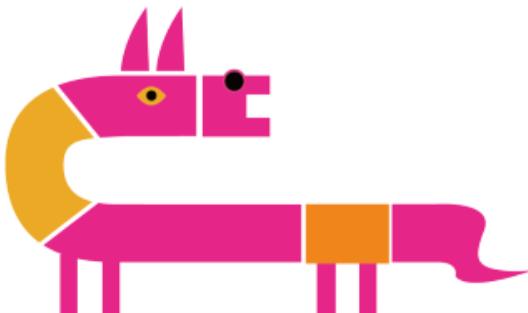
文件 地区、车站信息导入 天气信息导入 列车运行信息导入

从文件导入 从文件夹导入 上传缓存数据

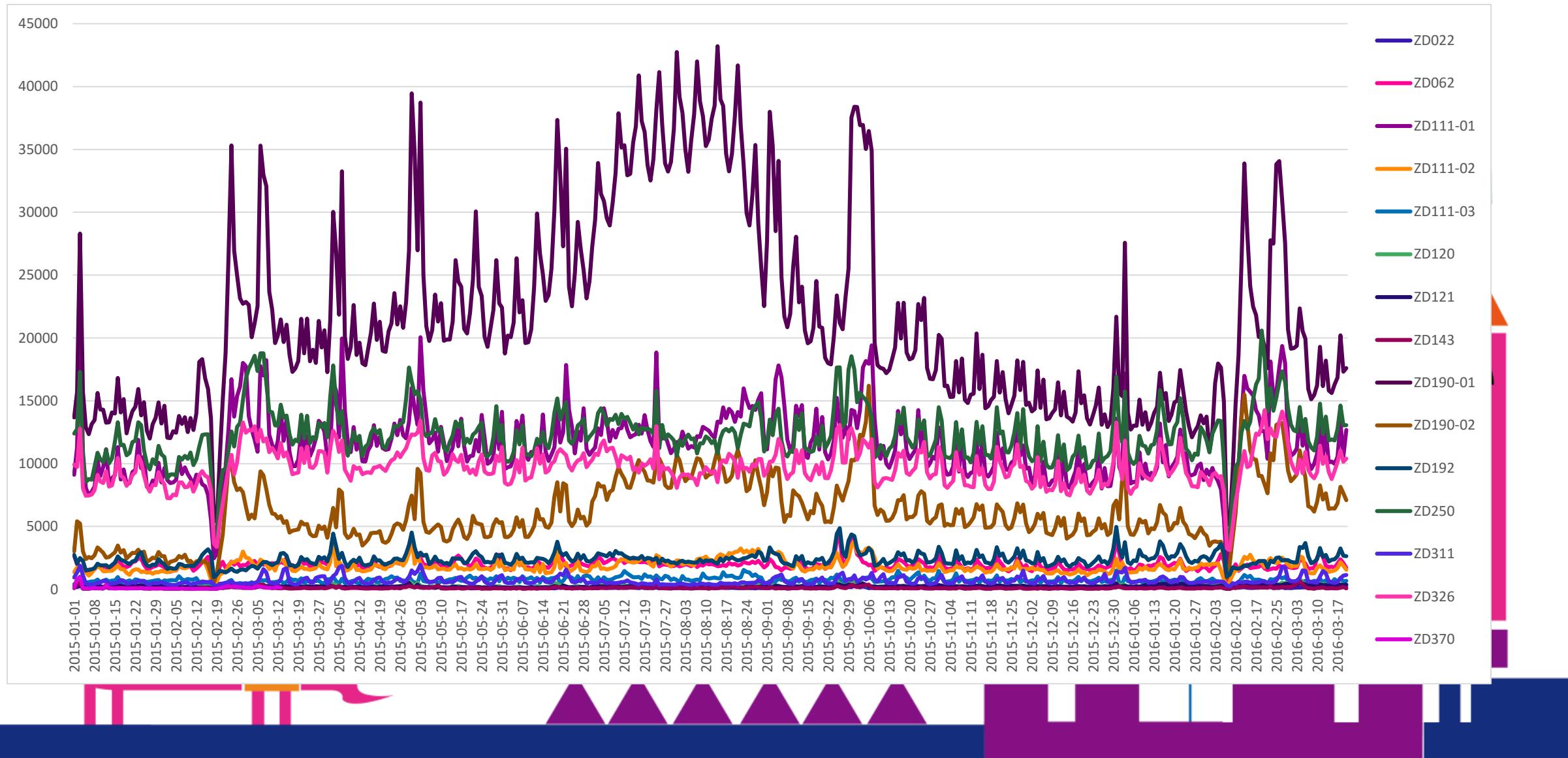
列车运行系统数据导入器

文件 地区、车站信息导入 天气信息导入 列车运行信息导入

[01:22:33]开始上传列车运行数据! 请勿关闭程序!
[01:22:33]预处理完成。
[01:22:33]正在处理2015-12-11的数据
[01:22:31]正在处理2015-12-10的数据
[01:22:29]正在处理2015-12-09的数据
[01:22:27]正在处理2015-12-08的数据
[01:22:24]正在处理2015-12-07的数据
[01:22:22]正在处理2015-12-06的数据
[01:22:21]正在处理2015-12-05的数据
[01:22:19]正在处理2015-12-04的数据
[01:22:17]正在预处理列车运行数据!
[01:22:17]开始导入列车运行数据! 请勿关闭程序!



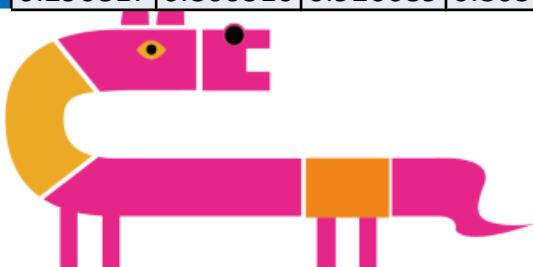
2015-01-01至2016-03-20 管内各车站客流量折线统计图



2015-01-01至2016-03-20

管内各车站客流量相关性系数表

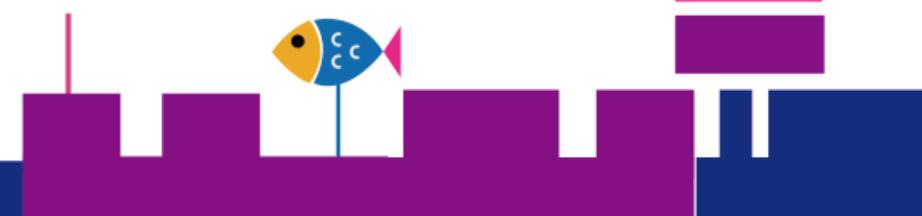
	ZD022	ZD062	ZD111-01	ZD111-02	ZD111-03	ZD120	ZD121	ZD143	ZD190-01	ZD190-02	ZD192	ZD250	ZD311	ZD326	ZD370	总客流
ZD022	1															
ZD062	0.472303	1														
ZD111-01	0.351596	0.738481	1													
ZD111-02	0.116033	0.771182	0.655772	1												
ZD111-03	0.060091	0.28769	0.633225	0.215463	1											
ZD120	0.269668	0.665113	0.554326	0.655175	0.244711	1										
ZD121	0.336353	0.294572	0.341477	0.239219	0.319976	0.627435	1									
ZD143	0.284482	0.601116	0.53582	0.580306	0.222649	0.807181	0.410456	1								
ZD190-01	0.079284	0.753366	0.708237	0.756161	0.285993	0.435881	-0.00121	0.391424	1							
ZD190-02	0.073614	0.691063	0.681622	0.696208	0.369117	0.671838	0.469365	0.513342	0.770668	1						
ZD192	0.442939	0.792535	0.730776	0.627082	0.504519	0.59658	0.371004	0.547269	0.542464	0.607591	1					
ZD250	0.437013	0.805823	0.946943	0.709193	0.571674	0.656202	0.436828	0.586202	0.676054	0.697468	0.78366	1				
ZD311	0.613909	0.625923	0.625641	0.376043	0.405073	0.487845	0.38095	0.492623	0.244547	0.296718	0.788901	0.676195	1			
ZD326	0.428367	0.781893	0.959185	0.675328	0.571268	0.595374	0.375006	0.553182	0.694667	0.674489	0.757008	0.9743	0.672511	1		
ZD370	0.237154	0.364958	0.387593	0.161195	0.191945	0.879408	0.396668	0.882661	0.230415	0.199881	0.416739	0.342154	0.575127	0.386495	1	
总客流	0.290817	0.860316	0.926689	0.803849	0.503903	0.633599	0.297836	0.570128	0.893502	0.834741	0.763452	0.923575	0.545252	0.925526	0.360622	1



管内车站客流量相关图表分析：

以上图表是根据15个月每天每个站的客流量所做的统计图及各站客流量的相关性系数表。

从客流量折线统计图中可以看出，ZD190-01在全年中客流量都稳居高位，特别是暑假期间客流量明显增多，且在法定节假日时急剧上升。作为当地地铁线路的最后一站，此站点客流量远远大于其他站点的客流量，对站点的承载力要求较高。此站点在当地是一个具有相当大影响力的站点，而当地存在其他客流量小的站点，可以做适当调整，以配合该站缓解客流压力。

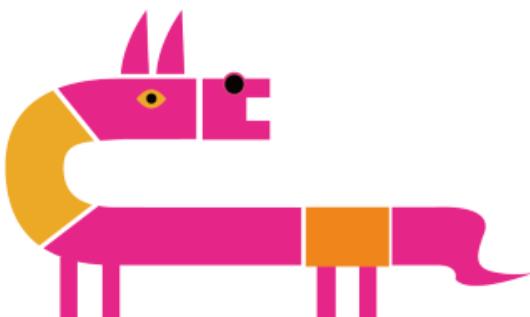


而ZD143、ZD311等站点在全年客流量都偏少，说明这些站点不处于城市中心地区，不作为当地主要车站，途径列车较少。此类车站使用潜能较大，可参照当地城市规划以及站点周边交通发展情况，增加停靠列车数量，缓解当地主要车站的客流压力，或者加开城际列车，发展城市快捷交通圈，以提高资源利用率。

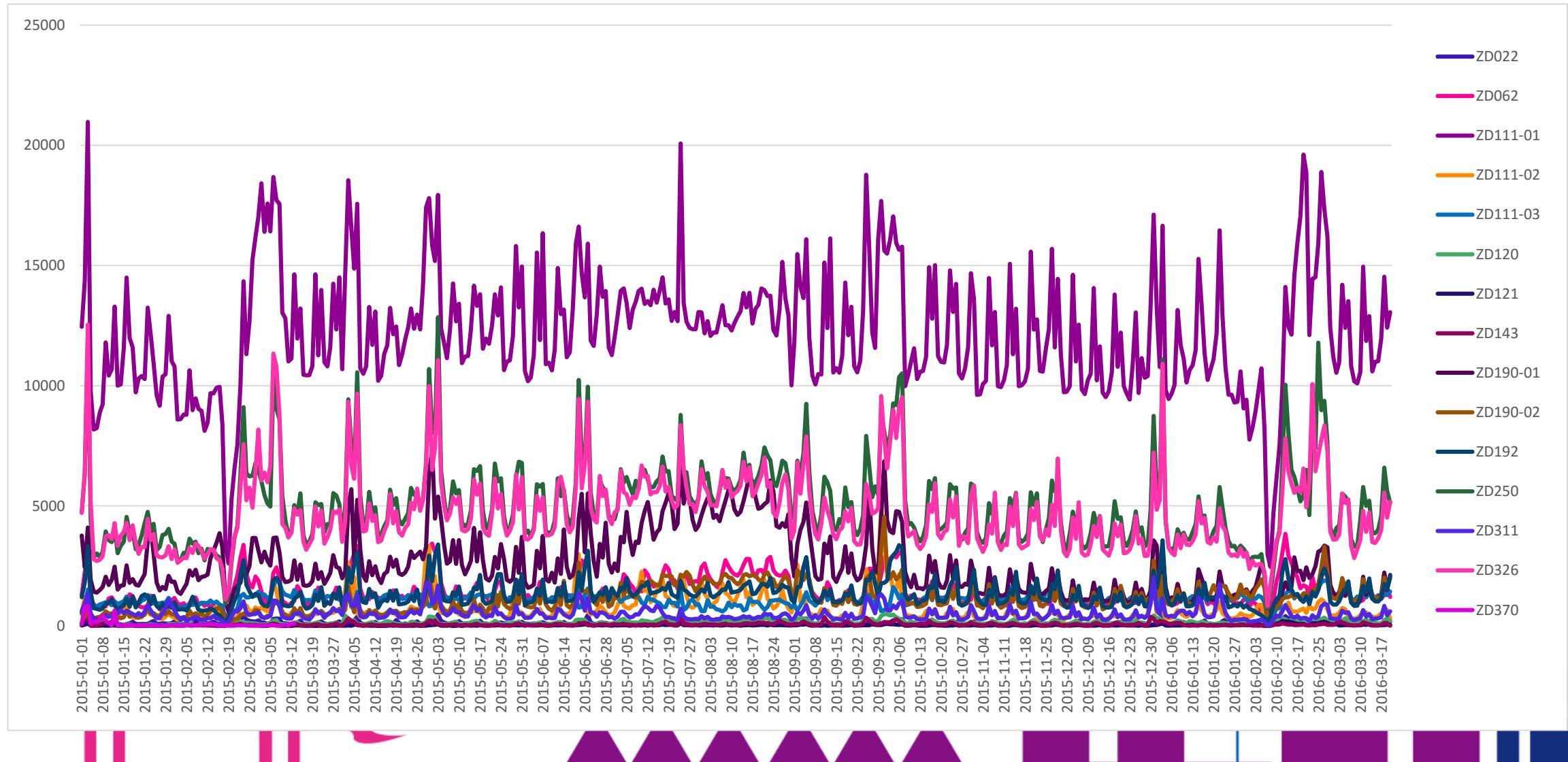
由客流量相关性系数表中可以看出，大部分车站之间的客流量呈中度正相关。说明管内车站的客流量规律有一定的相同性，当管内某一个车站出现大幅度客流变化的时候，应通知其他车站做好相关工作，防止客流压力突增带来的相关问题。



而ZD111-01、ZD250、ZD326这三个车站两两之间的相关系数大于0.95，说明这三个车站的客流日客流量曲线呈高度正相关。说明这三个站在管内铁路线上的地位相当，当一个车站出现客流变化时，另外两个车站很有可能出现相同情况。为此这三个车站可以形成协议，讨论出共同的方案，保证各站能在最短时间预防客流量的变化带来的问题。



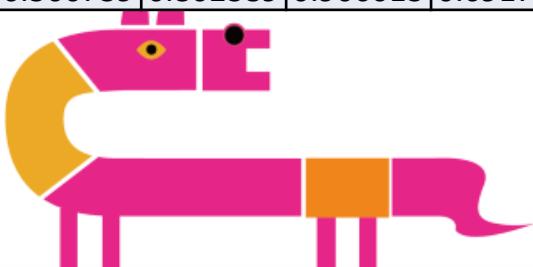
2015-01-01至2016-03-20期间 管内各车站上车人数折线统计图



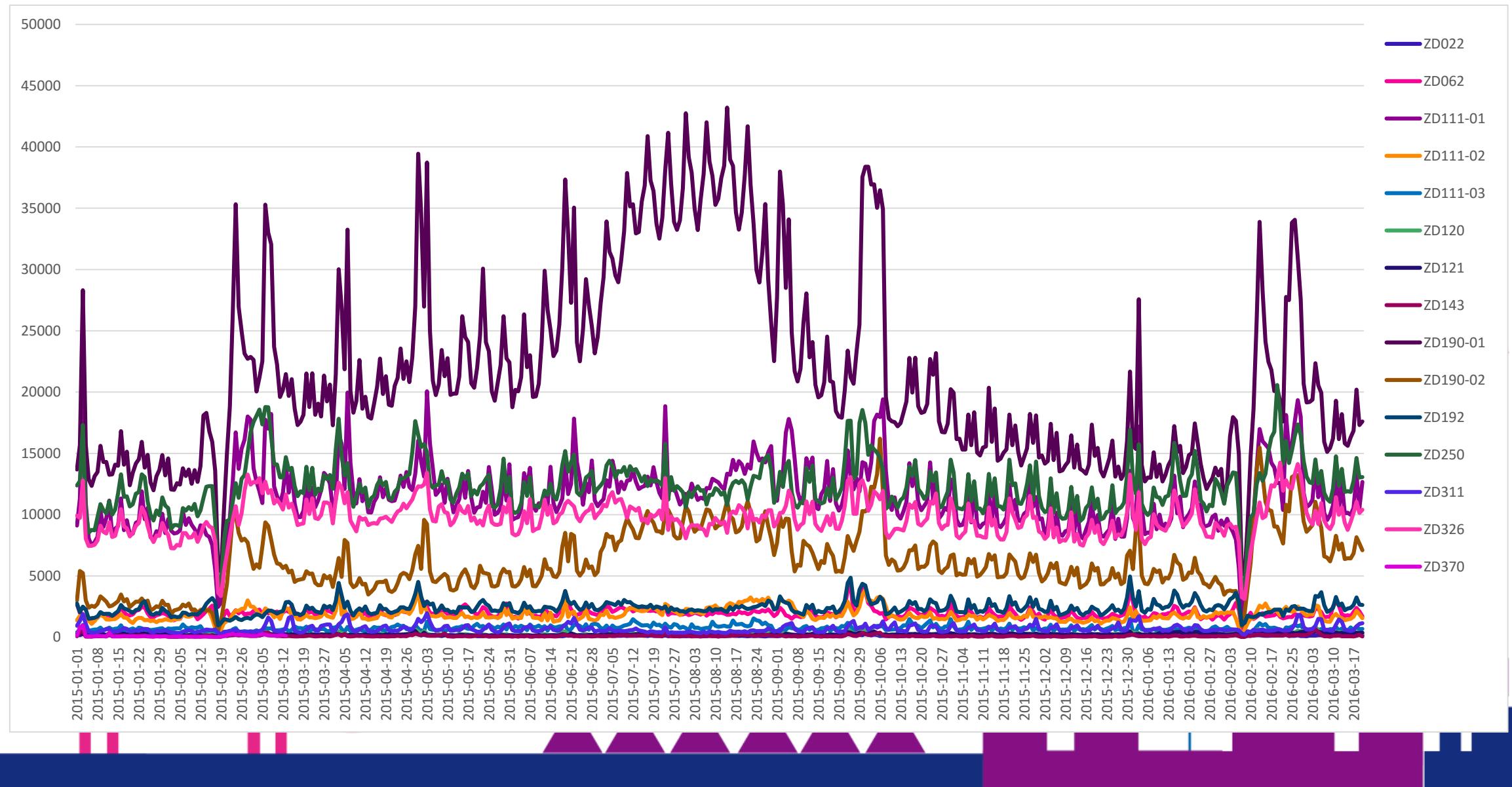
2015-01-01至2016-03-20

管内各车站上车人数相关性系数表

	ZD022	ZD062	ZD111-01	ZD111-02	ZD111-03	ZD120	ZD121	ZD143	ZD190-01	ZD190-02	ZD192	ZD250	ZD311	ZD326	ZD370	总人数
ZD022	1															
ZD062	0.671419	1														
ZD111-01	0.429134	0.554961	1													
ZD111-02	0.108238	0.538603	0.56646	1												
ZD111-03	0.343945	0.067235	0.535744	-0.08789	1											
ZD120	0.305768	0.644657	0.521707	0.612731	0.073702	1										
ZD121	0.554107	0.548302	0.278595	0.198641	0.226418	0.506093	1									
ZD143	0.201197	0.424034	0.451385	0.43848	0.122829	0.726942	0.237714	1								
ZD190-01	0.299257	0.684027	0.595007	0.608384	-0.04515	0.458841	0.155143	0.338213	1							
ZD190-02	0.172565	0.527637	0.500423	0.578592	0.105205	0.612826	0.408826	0.371151	0.594265	1						
ZD192	0.709586	0.811205	0.681072	0.466298	0.351629	0.580189	0.477076	0.49957	0.55309	0.538243	1					
ZD250	0.639209	0.822616	0.827093	0.608956	0.381151	0.62166	0.47918	0.490894	0.681527	0.566032	0.870056	1				
ZD311	0.426445	0.428682	0.654092	0.52871	0.370156	0.427846	0.20392	0.440088	0.353183	0.308003	0.630283	0.660521	1			
ZD326	0.631287	0.790548	0.826838	0.62703	0.347463	0.594161	0.438134	0.445664	0.673769	0.528923	0.838202	0.953241	0.635149	1		
ZD370	0.20309	0.232662	0.22827	0.129491	0.01567	0.829032	0.436584	0.876488	0.123899	-0.00244	0.348654	0.27249	0.309059	0.284699	1	
总人数	0.566739	0.802389	0.906013	0.691756	0.366039	0.657264	0.419807	0.516598	0.769972	0.646644	0.847534	0.963023	0.675448	0.953515	0.291771	1



2015-01-01至2016-03-20期间 管内各车站下车人数折线统计图



2015-01-01至2016-03-20

管内各车站下车人数相关性系数表

	ZD022	ZD062	ZD111-01	ZD111-02	ZD111-03	ZD120	ZD121	ZD143	ZD190-01	ZD190-02	ZD192	ZD250	ZD311	ZD326	ZD370	总计
ZD022	1															
ZD062	0.398033	1														
ZD111-01	-0.05559	0.384596	1													
ZD111-02	-0.05397	0.520911	0.696287	1												
ZD111-03	-0.29298	0.278972	0.492514	0.410109	1											
ZD120	0.129949	0.5125	0.434846	0.499549	0.161644	1										
ZD121	0.182911	0.239078	0.239751	0.183812	0.053124	0.660679	1									
ZD143	0.15399	0.415283	0.460236	0.4906	0.095029	0.753114	0.432362	1								
ZD190-01	-0.21866	0.397786	0.705825	0.700897	0.465703	0.324506	-0.08491	0.376944	1							
ZD190-02	-0.26792	0.270576	0.698391	0.665541	0.386016	0.585091	0.37074	0.505988	0.778228	1						
ZD192	0.305086	0.780556	0.366272	0.455125	0.34932	0.472592	0.30873	0.326236	0.293652	0.289038	1					
ZD250	0.201199	0.653058	0.741829	0.621574	0.349991	0.626692	0.448589	0.532563	0.499993	0.557521	0.60628	1				
ZD311	0.363747	0.371605	0.47418	0.182382	0.060564	0.333696	0.274186	0.326573	0.111542	0.21798	0.519476	0.479321	1			
ZD326	0.149048	0.595214	0.800857	0.578438	0.382226	0.525124	0.382193	0.450526	0.506342	0.540466	0.554264	0.928429	0.502746	1		
ZD370	-0.0678	0.128897	0.520818	0.15123	0.145104	0.88609	0.398354	0.819363	0.404162	0.418999	-0.02068	0.372949	0.638332	0.380266	1	
总计	-0.09114	0.542413	0.879543	0.792817	0.510207	0.534612	0.195367	0.522901	0.916589	0.850644	0.481097	0.771339	0.343877	0.773372	0.465214	1

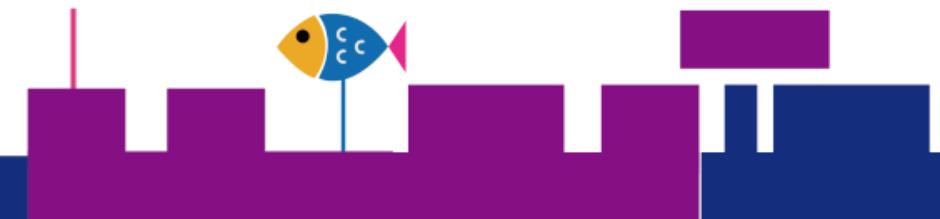
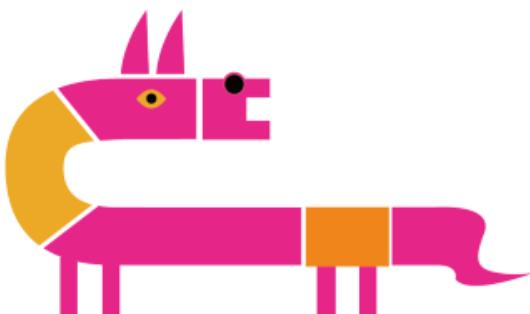


管内车站上下车人数图表分析：

以上图表分别是根据15个月每天每个站的上下车人数所做的统计图及各站间每日上车人数曲线与下车人数曲线的相关性系数表。

从各车站下车人数折线统计图中可以看出，ZD190-01等站点7月到10月上车人数普遍较高，说明此车站流出人口较多，是人流高峰期，可能是旅游专区，预留车票数量可以适当增加。

在各车站下车人数折线统计图中我们可以明显看出，ZD111-01、ZD190-01等站点对客流有较大吸引力度。

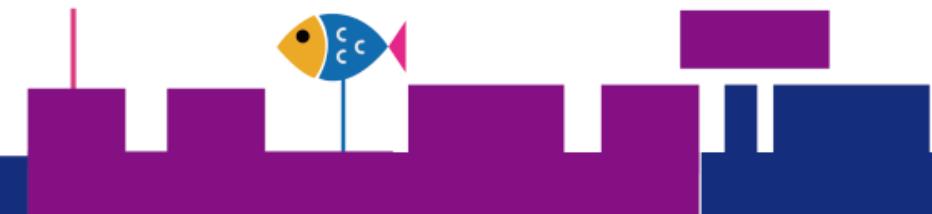


同时我们也可以得出推论：这些站点的流动人口数、城市化程度以及经济发展水平相对比其他站点高。

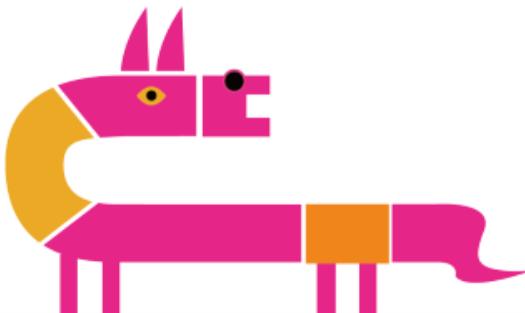
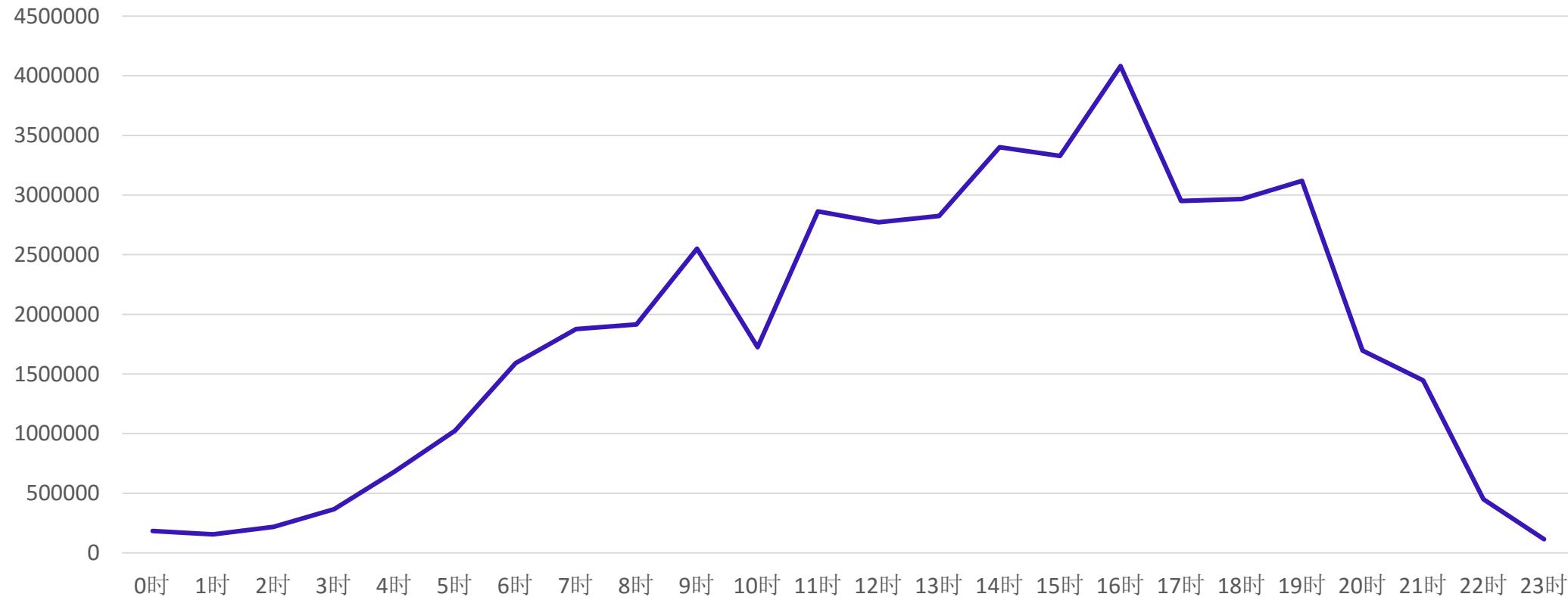
ZD143、ZD311等站点在全年上下车人数都偏少，说明此站点客流量不高，车站资源利用率不高。

对比两张图可以发现，管内车站的下车人数明显大于。可以发现客流量有一部分是从管外车站流动到管内车站。

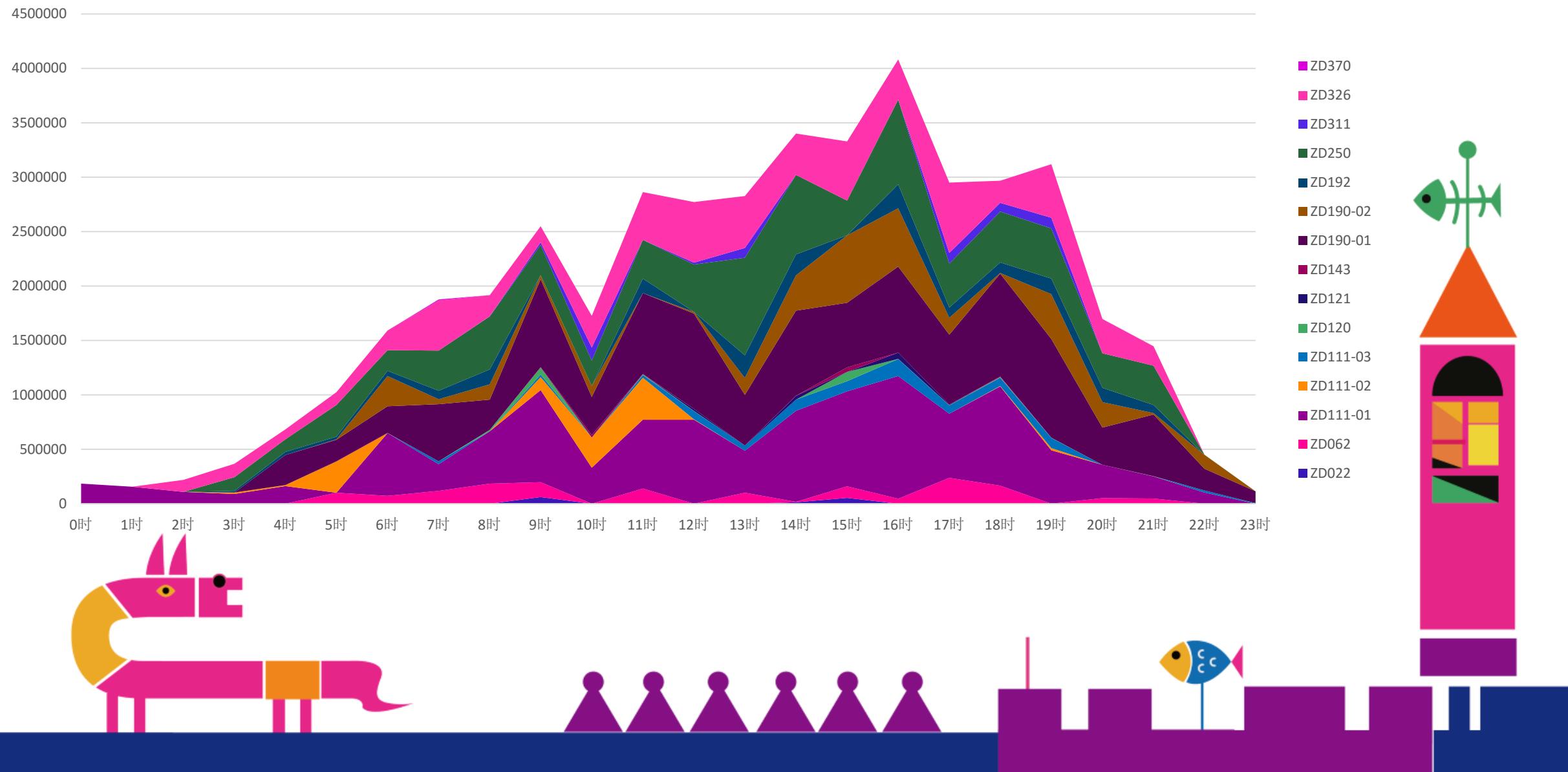
管内各车站的上下车人数的相关性系数明显低于上车量。说明各站点间的客流流动不平均，尽管各车站客流量有相同规律，但是各站还是需要单独拿出方案应对乘车人流与抵达人流的变化。



2015-01-01至2016-03-20
管内车站每小时总客流量折线统计图



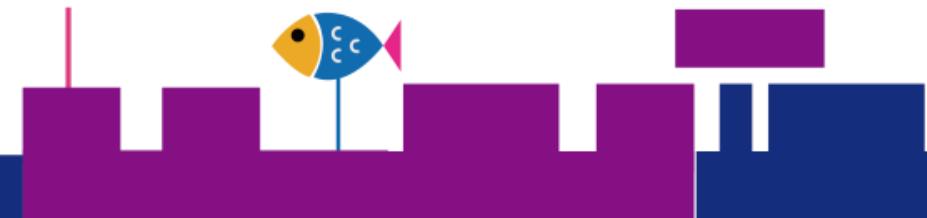
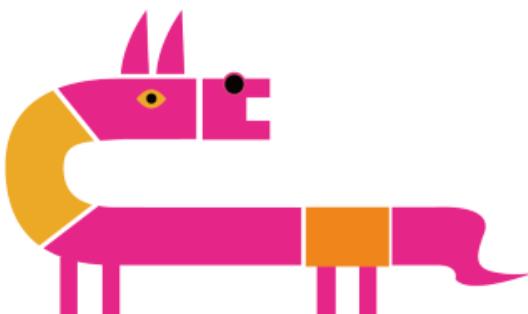
2015-01-01至2016-03-20 管内各车站每小时客流量面积图



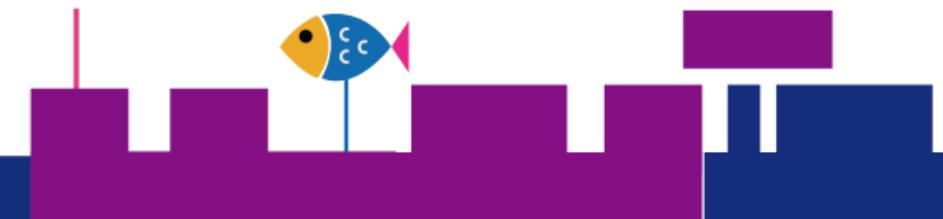
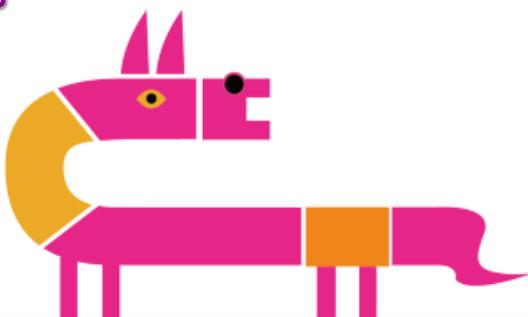
管内车站每小时客流量图表分析：

由本图按照时间序列分析一天中管内15个车站总客流变化趋势，反映出一天内的客流量显示出较强的不均衡性。本折线统计图还可以反映出一定的规律性：凌晨时间客流量较低，各个站点上午的高峰小时并不明显，从6:00-7:00开始客流量逐渐增多，16:00到达峰值，之后缓慢下降。

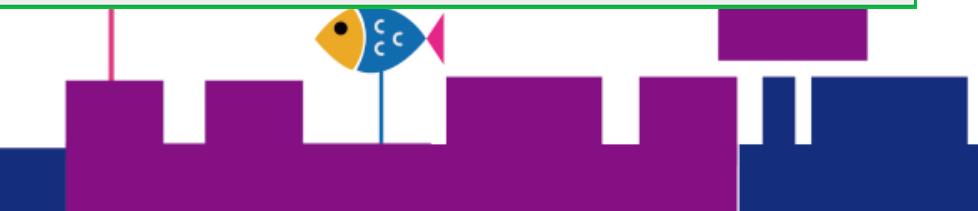
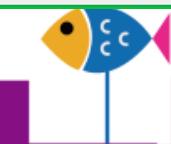
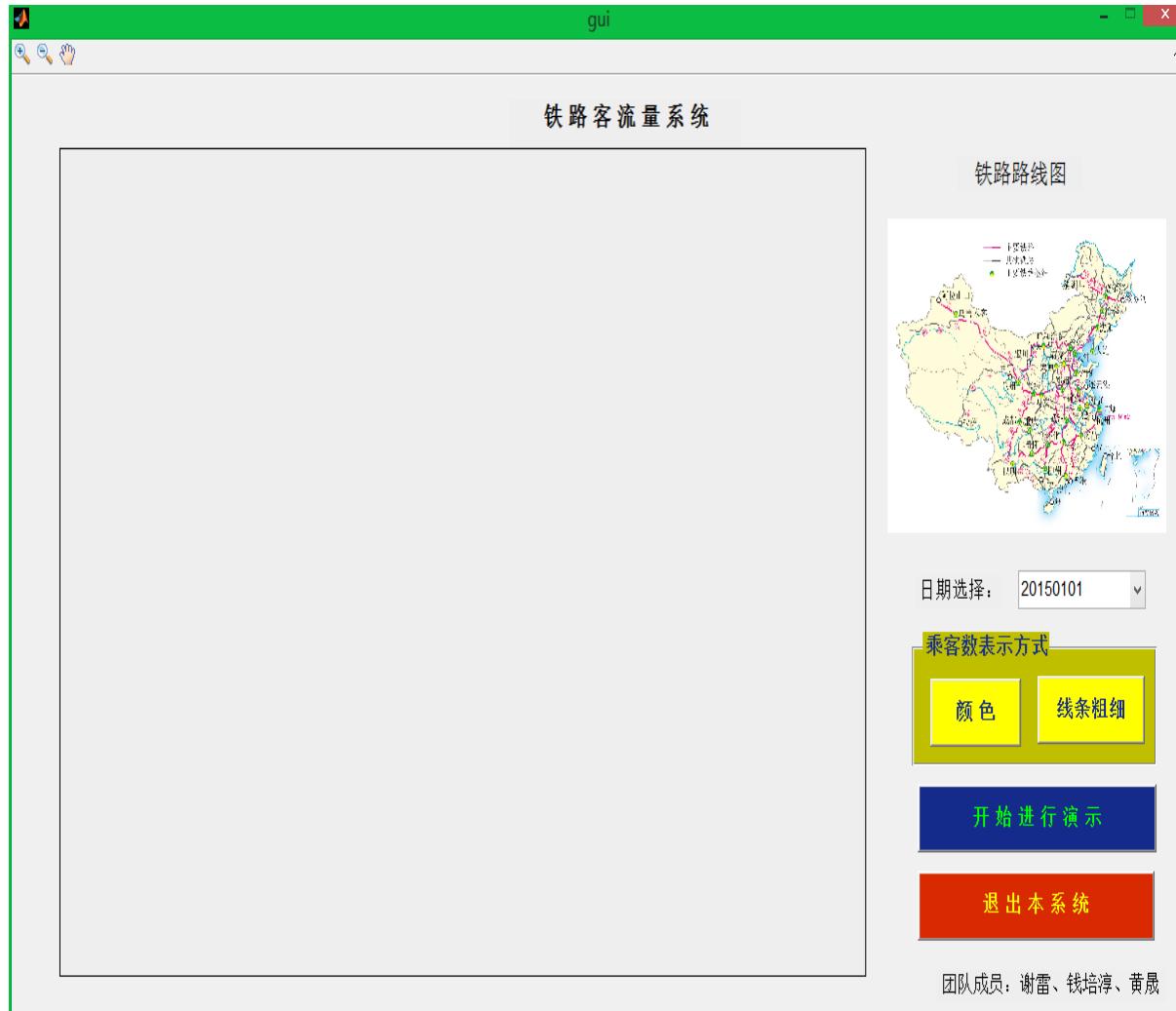
人口流动在白天相对频繁，大多数旅客选择白天搭乘列车，具有多方面原因，分析如下：

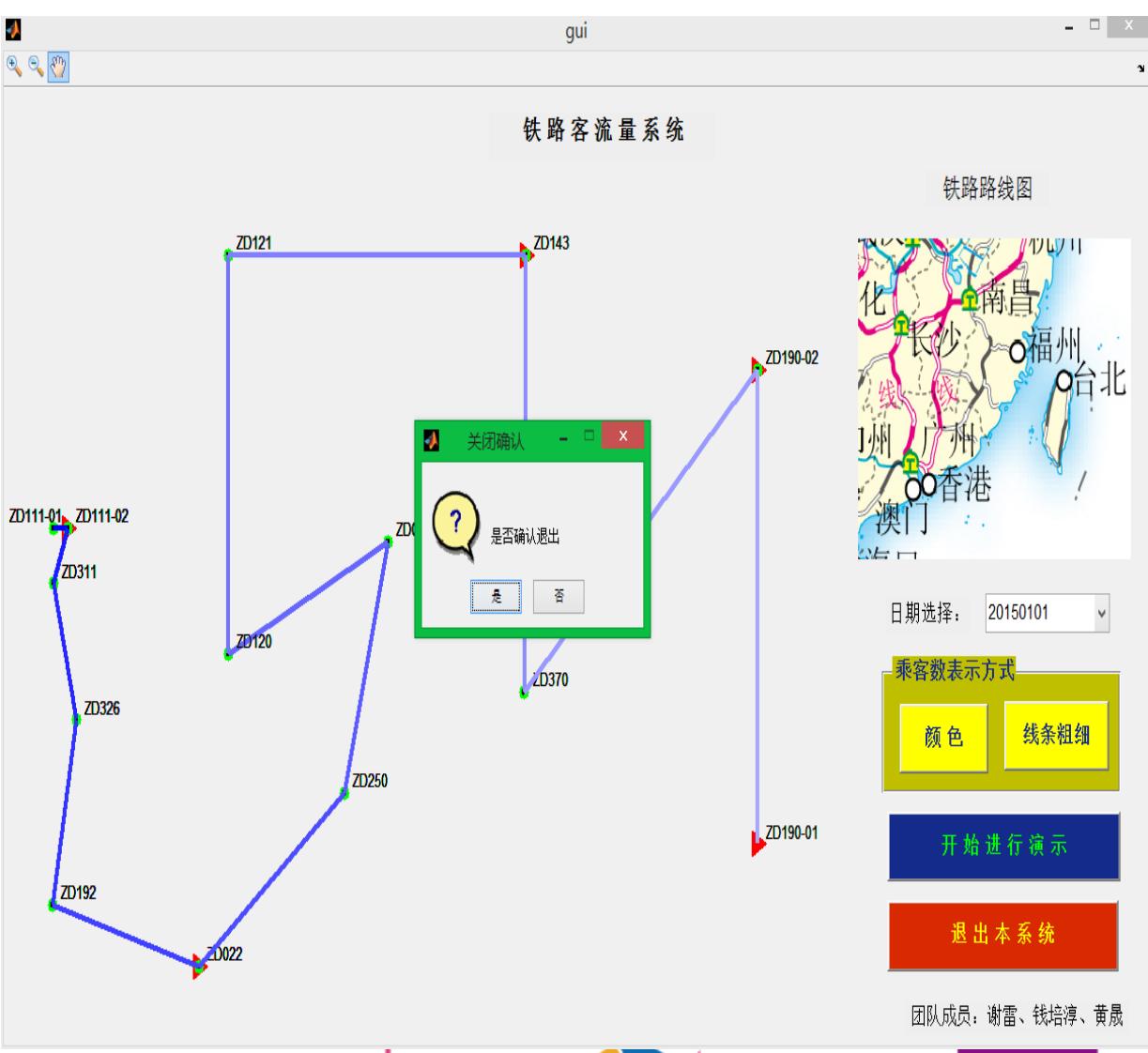
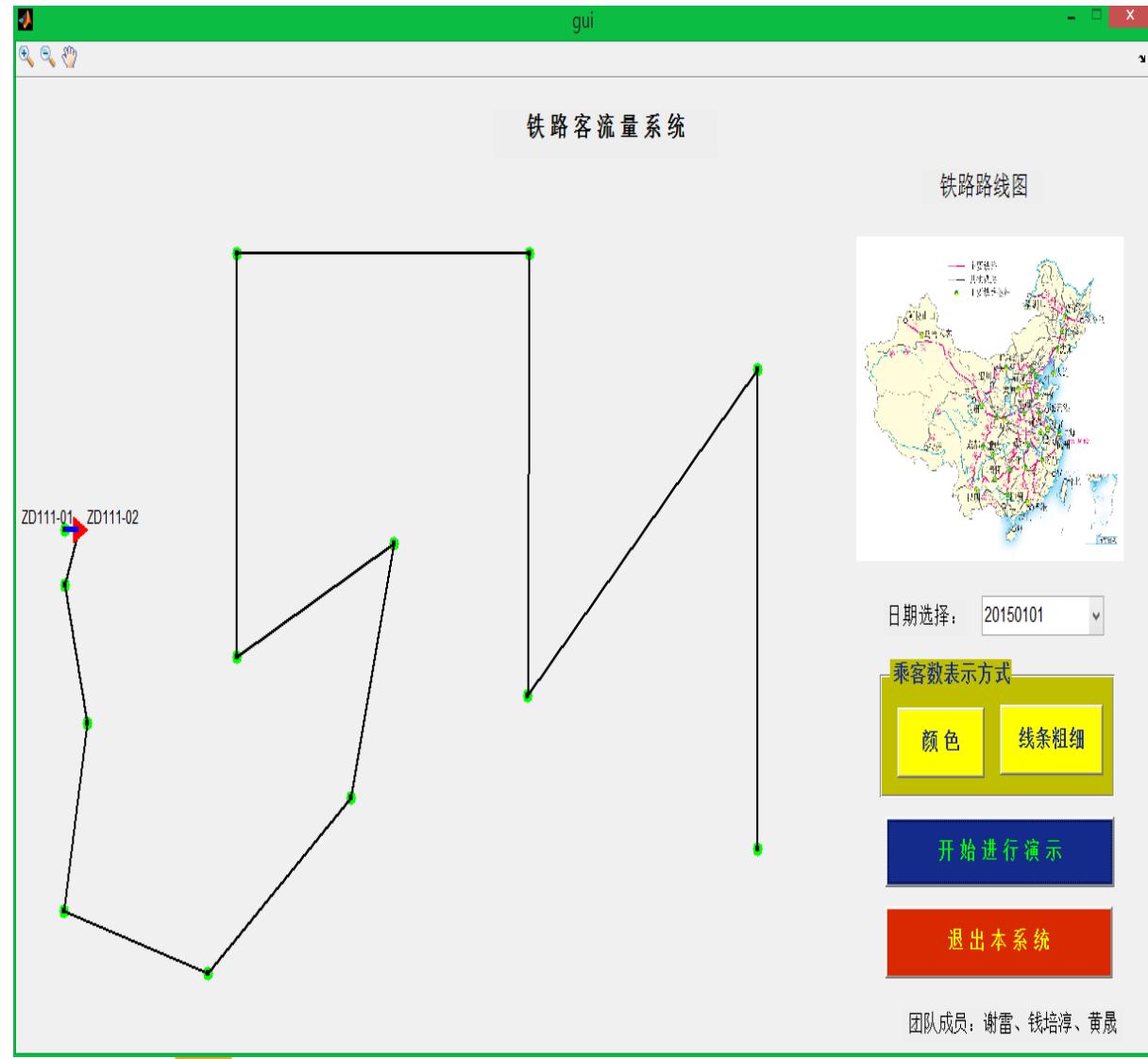


不同地区经济发展水平不同，存在梯度性差异，因此不同地区的旅客总体消费水平也存在着较大差异。一天内管内所有车站总客流变化主要由各个人口密度较大、经济水平较高的地区所决定。平日里的客流以工人、企业人员和学生为主，主要出行目的是公务、会议、商务、务工、学习培训、探亲、旅游休闲等，都聚集在白天几个主要时间段。在春运、五一、十一等客运旺季时，客流主要还是学生流、探亲流和商务流等，旅客基本都会选择白天出行。充分体现了如今人们生活节奏快，对工作和旅行的需求。



列车客流量系统：





系统描述：

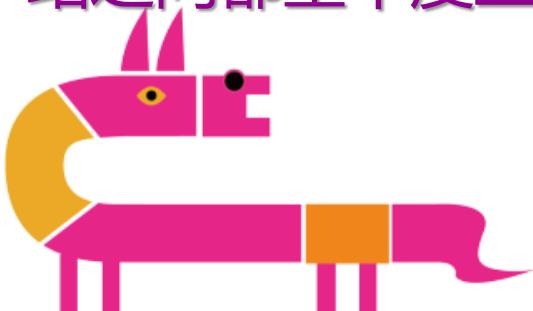
本系统提供了中国铁路路线图，可使用菜单栏中的缩放，移动功能进行相关查看。日期选择中有选择菜单，可针对不同日期进行站点客流演示。本系统的设计针对用户不同需求提供了乘客数表示方式：颜色、线条粗细，可以更为直观地查看。

本系统主要功能是在一定时间内对铁路公司管内各站点乘客数变化进行动态演示。说明（红色箭头为列车）：列车从站点出发时会有乘客上车，抵达站点时有乘客上车也有乘客下车。在经过一个站点时，其乘客数就可能会发生改变。



总结及建议：

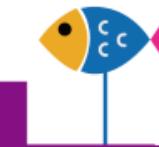
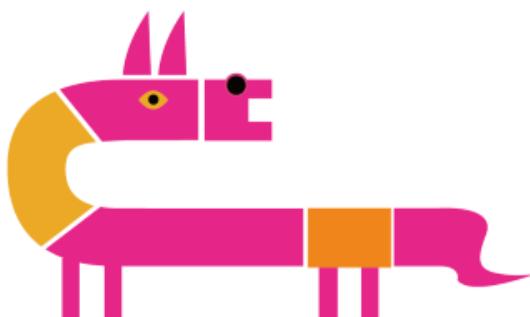
- 1、该铁路沿线各站的客流有不同特点，在一天内有各自明显的出行规律，大部分站点在节假日期间都会形成一定的高峰时段，旅客流动相对频繁。
- 2、特别是针对ZD190-01等上车人数较多的站点我们应该提高预留车票数量。
- 3、从客流量、上下车人数相关性系数表中可以看出，大部分车站之间都呈中度正相关。



4、针对不同的时间段，建议在旅客流量较大的站点处加派安保人员，疏导人群，避免安全事故的发生。

5、对于ZD143等小站点，可以减少车次以及合理安排车辆配置。

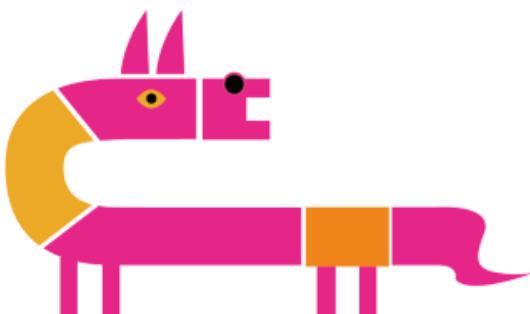
6、列车通过停站方式输送旅客，停站数目越多越有利于旅客的乘降，方便旅客出行。但列车中途停站过多会降低旅行速度，所以列车停站的设置既要考虑旅客出行的方便性，又不能过多地降低列车旅行速度，从而造成旅客旅行时间的增加。



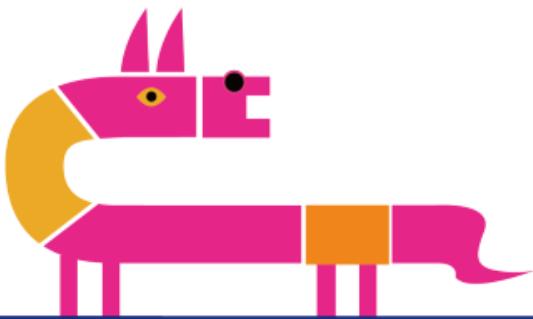
在此给出几个停站方案（适用于高速列车）：

- ① 仅在指定的个别客运站始发、终到，且在运行径路上除始发、终到站以外的任何客运站都不办理客运作业。
- ② 在选定的客运站上办理客运作业。

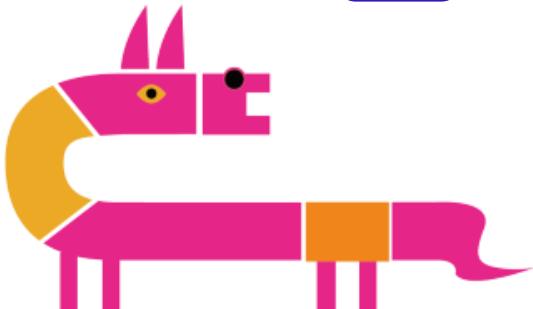
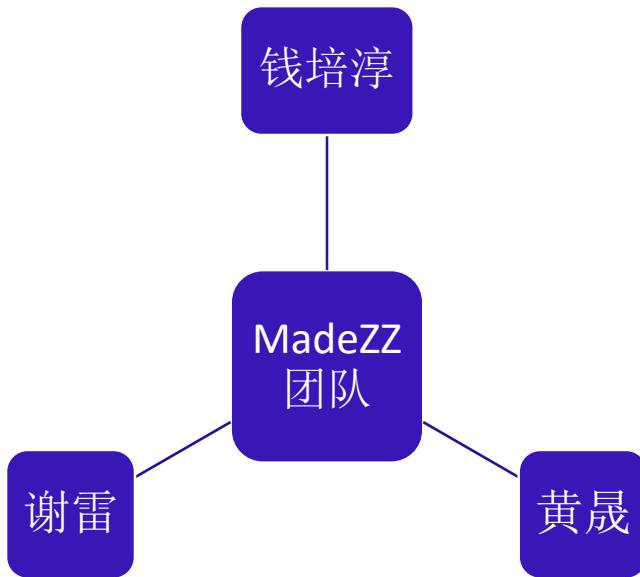
7、对于客座率超过100%的车次，说明需要售出站票。建议提高发车密度，以提高铁路客运部门提供给旅客的运输服务产品的数量水平。



8、合理公道的票价是旅客选择出行方式的一个重要因素。对于我国而言，铁路票价对旅客的影响具有一定作用。消费水平不同的旅客对于“合理的票价”有着不同的衡量标准。建议铁路部门可以更多考虑票价的合理性和旅客的承受能力来确定灵活的、动态的票价。



THANKS



MadeZZ团队
2016-05-09

