

大数据助力 民航智慧出行

中国民航信息网络股份有限公司

彭明田

11/16/2018



目录



1

民航进入大数据时代

2

民航旅客服务智能化实验室

3

实验室应用

4

互动环节(Q&A)



目录



1

民航进入大数据时代

2

民航旅客服务智能化实验室

3

实验室应用

4

互动环节(Q&A)

从行业视角来看

- 航空业的营业额增长幅度是全世界人均生产总值增长幅度的近**3倍**
- 全世界平均每年有**60亿**张机票售出，平均每天有**93000**次航班
- 至少**1千万吨CO2**由飞机排放在大气中



60亿张机票



93000次航班



1千万吨CO2

从民航旅客服务数据存储的视角来看

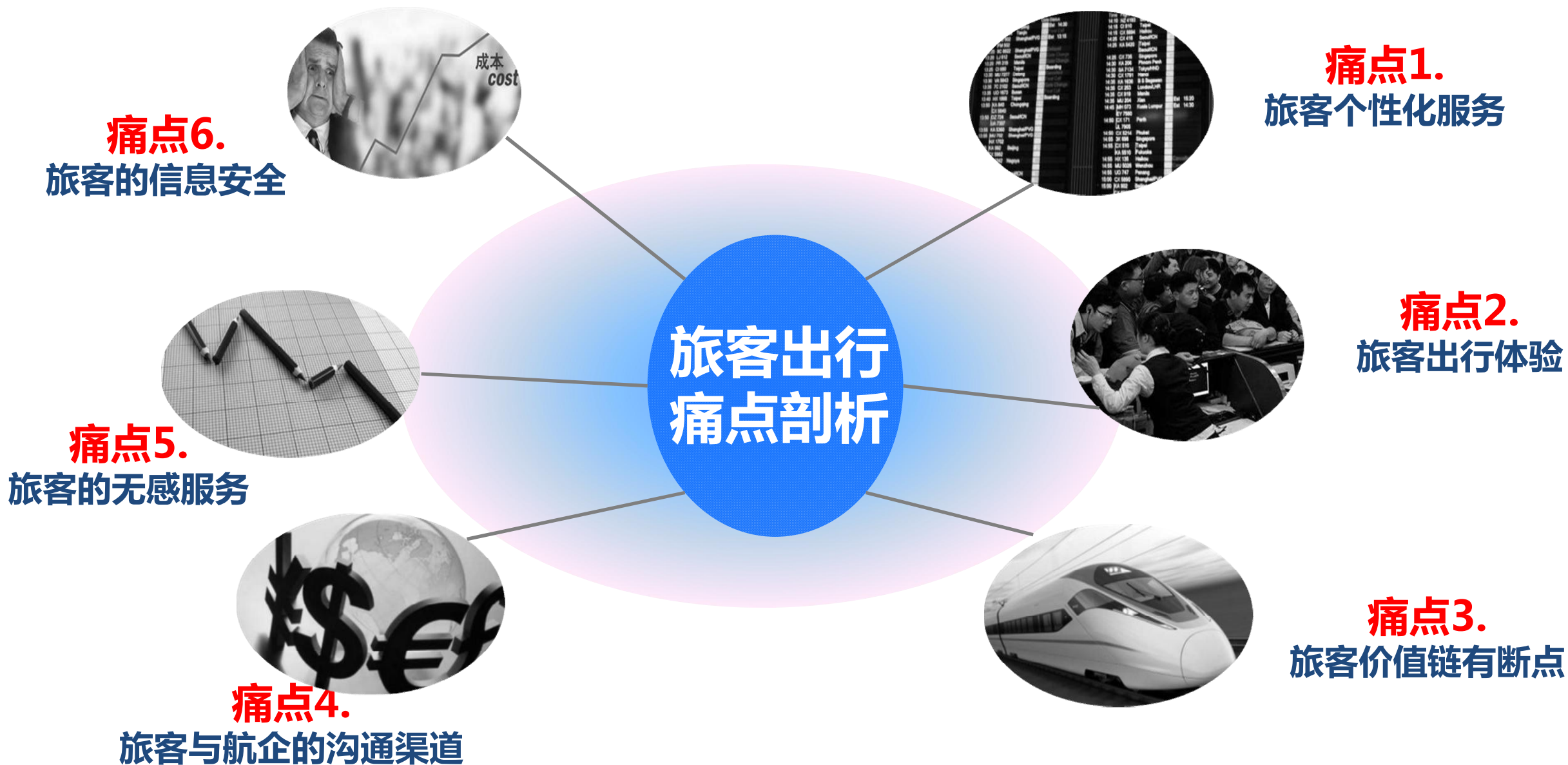


数据源自航信核心业务交易系统——旅客信息服务系统（PSS），PSS系统运营了38家航空公司的信息服务，业务范围涵盖航空产品管理、航空产品销售、机场旅客服务等领域，**每日**产生**20G+**的旅客交易相关数据、**5T+**的系统处理相关数据，**每年数据增量 1.5PB+**。

数据规模



5年间，伴随航信业务量增长与数字化系统建设的技术发展，数据量呈激增态势，数据规模扩充了**600余倍**





目录



1

民航进入大数据时代

2

民航旅客服务智能化实验室

3

实验室应用

4

互动环节(Q&A)



民航旅客服务智能化应用技术重点实验室

Key Laboratory of Intelligent Passenger Service of Civil Aviation



建设民航旅客服务智能化应用技术重点实验室
以实现民航旅客出行的智能化、个性化为目标，
从而为综合交通运输规划提供决策依据，提升全
行业的旅客服务水平。

创新



智慧



智能

民航旅客服务智能化应用技术实验室是以大数据科技平台为依托，利用人工智能、机器学习等技术实现民航旅客智慧出行，全面提升行业旅客服务水平。



实验室价值体系

1 高价值数据整合

规划 按业务场景完成数据整合

过去 基础数据提供（无目标）

平台能力：SaaS 化数据交付

工具运维

2 业务洞察/智能决策支持

规划 产品深度融合·用户粘性佳

过去 即时性不强且阅过即焚

服务能力：社区化协同开发

3 数据驱动业务创新

规划 数据产品规模化量产
数据科学技术工程化

过去 业务挖掘深度及产能不足

服务能力：数据挖掘分析
思维启蒙/方法培训

4 跨业合作/生态链共赢

规划 运营业务+数据+用户

过去 运营系统

服务能力：数据产品运营（咨询）

低

服务输出

业务价值度

产品运营

高



科技平台总体架构

中国航信
TravelSky

大数据应用过程

数据应用

数据挖掘

数据整合

数据存储

数据处理

数据采集

大数据处理模式

批量数据提供

接口提供

异构数据库同步

数据整合及分析应用

数据挖掘及机器学习

离线数据计算模式

流式数据计算模式

大数据架构

智能化支持

CEP引擎

实时检索

WEBSERVICE

OLAP

大数据挖掘

Machine Learning

数据挖掘算法库

大数据可视化平台
(Tableau)

大数据分析

离线分析

实时分析

流式计算

数据开发

大数据存储

Hdfs

Redis

Greenplum

Hbase

Impala

EDB

.....

大数据计算引擎

MapReduce

检索引擎

SQL

Spark

Storm

MPI

.....

HDFS

Kafka集群

ETL工具

FtpOverHdfs

数据采集

消息采集

平台管理

任务调度

元数据

数据质量

数据安全

平台计量

资源负载

监控告警

■ 航信的自有数据

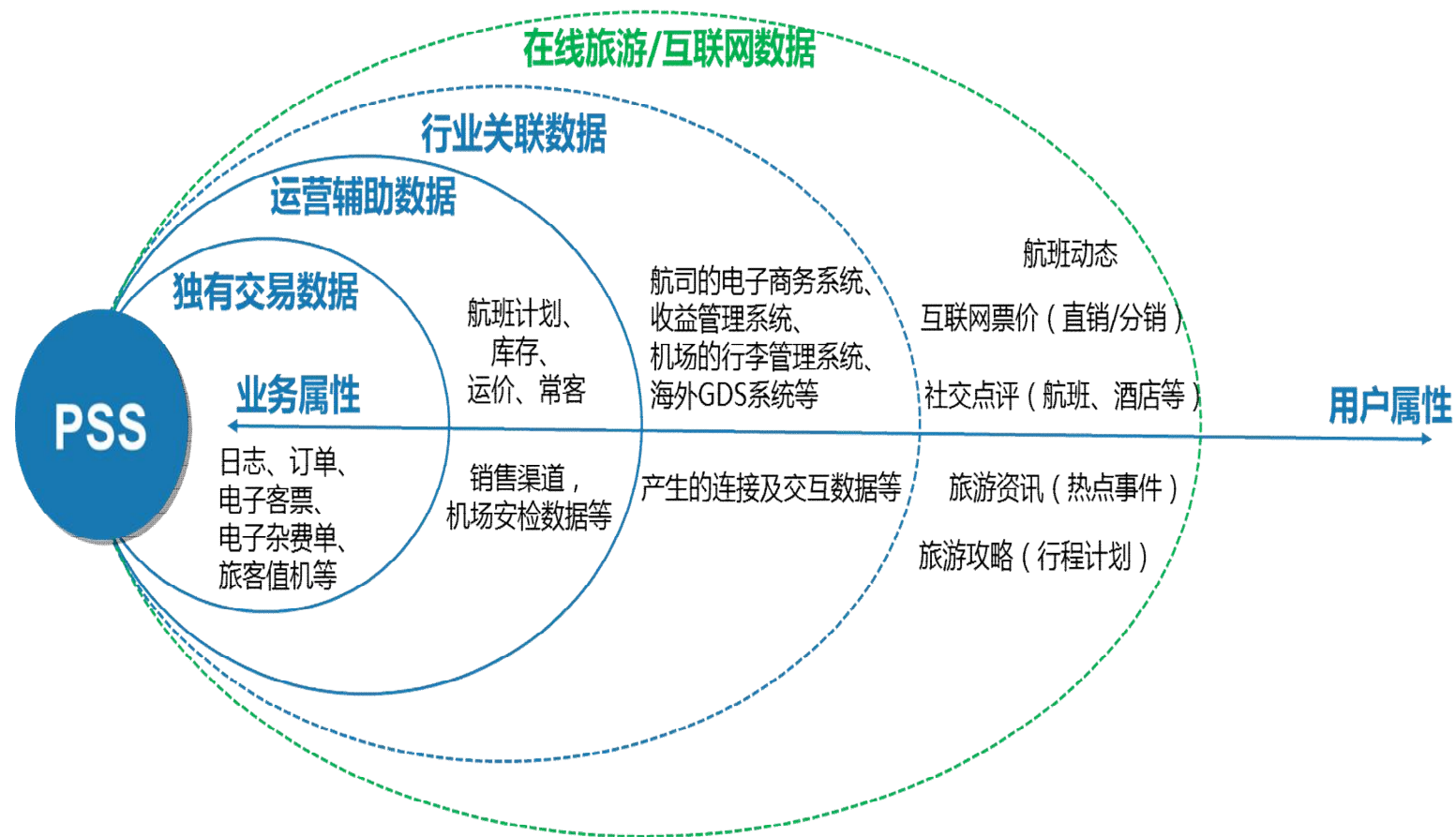
PSS业务主体产生

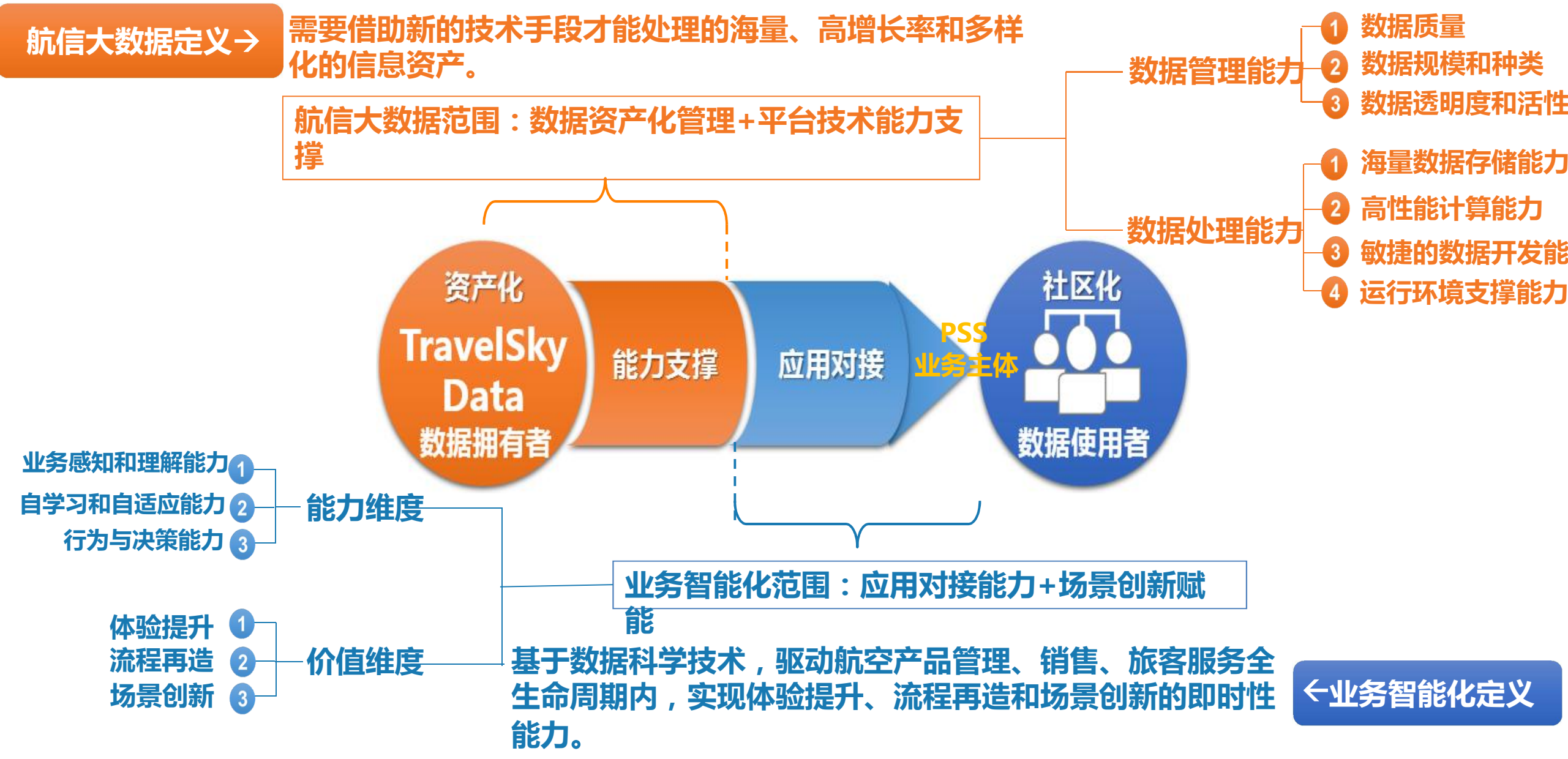
独有交易数据+运营辅助数据为主

■ 可获取的外部数据

行业合作伙伴链接及交互数据，

互联网数据（航班/运价/酒店/点评）



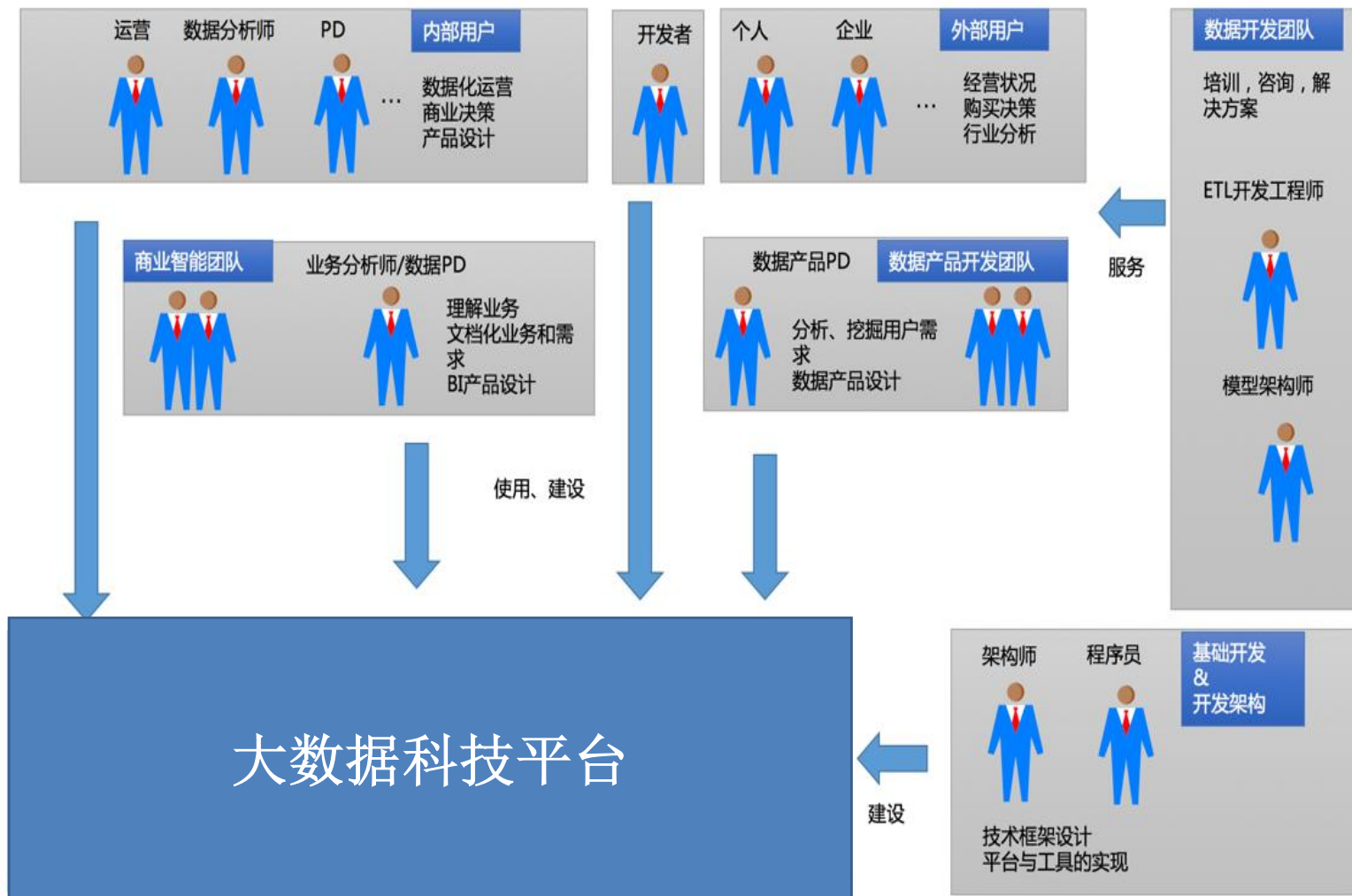




- 目前与院校合作
(北京交通大学、中国民航大学) 共同建设实验室



- 未来将继续科技平台建设，
建立开放化社区，成为民航
科技创新的孵化器





目录



1

民航进入大数据时代

2

民航旅客服务智能化实验室

3

实验室应用

4

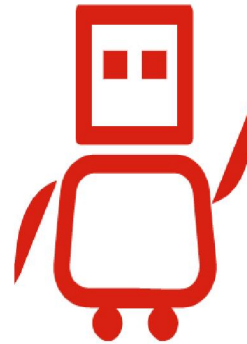
互动环节(Q&A)



两个平台案例



旅客画像



民航大脑



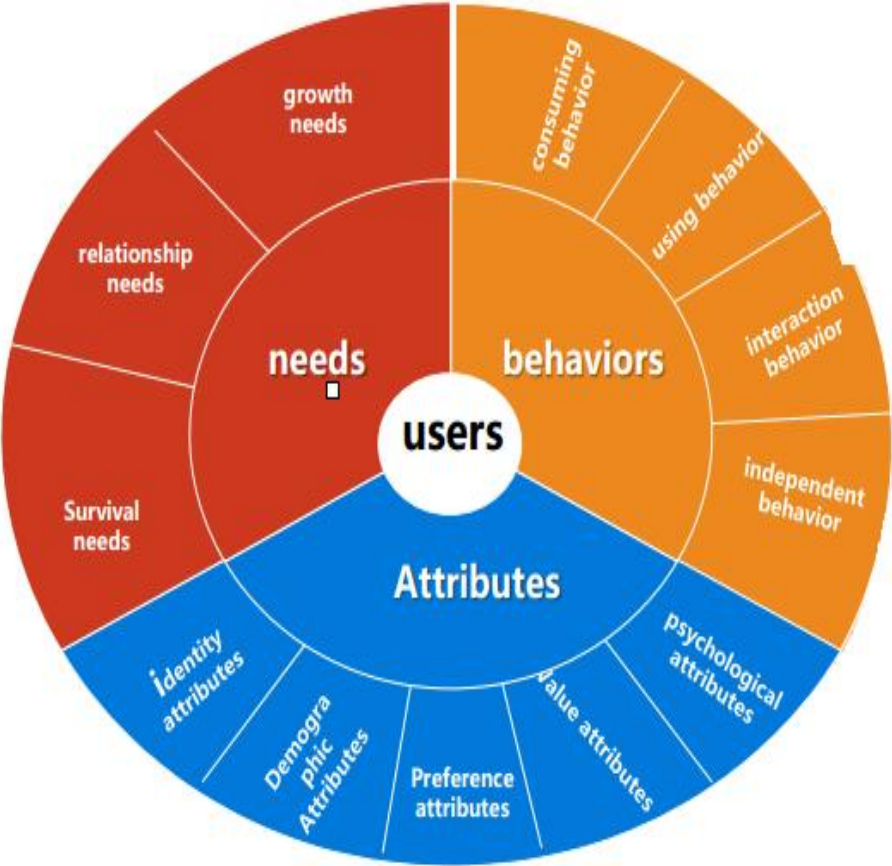
1.旅客画像



 购物画像	 航班画像	 设备画像	 质量保障	 旅客画像
 代理人画像	 航线画像	 景点画像	 商城画像	 城市画像



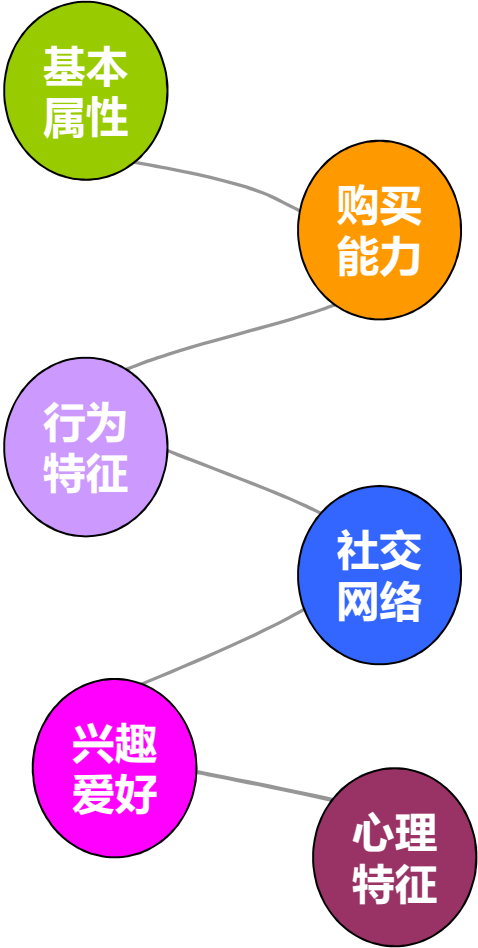
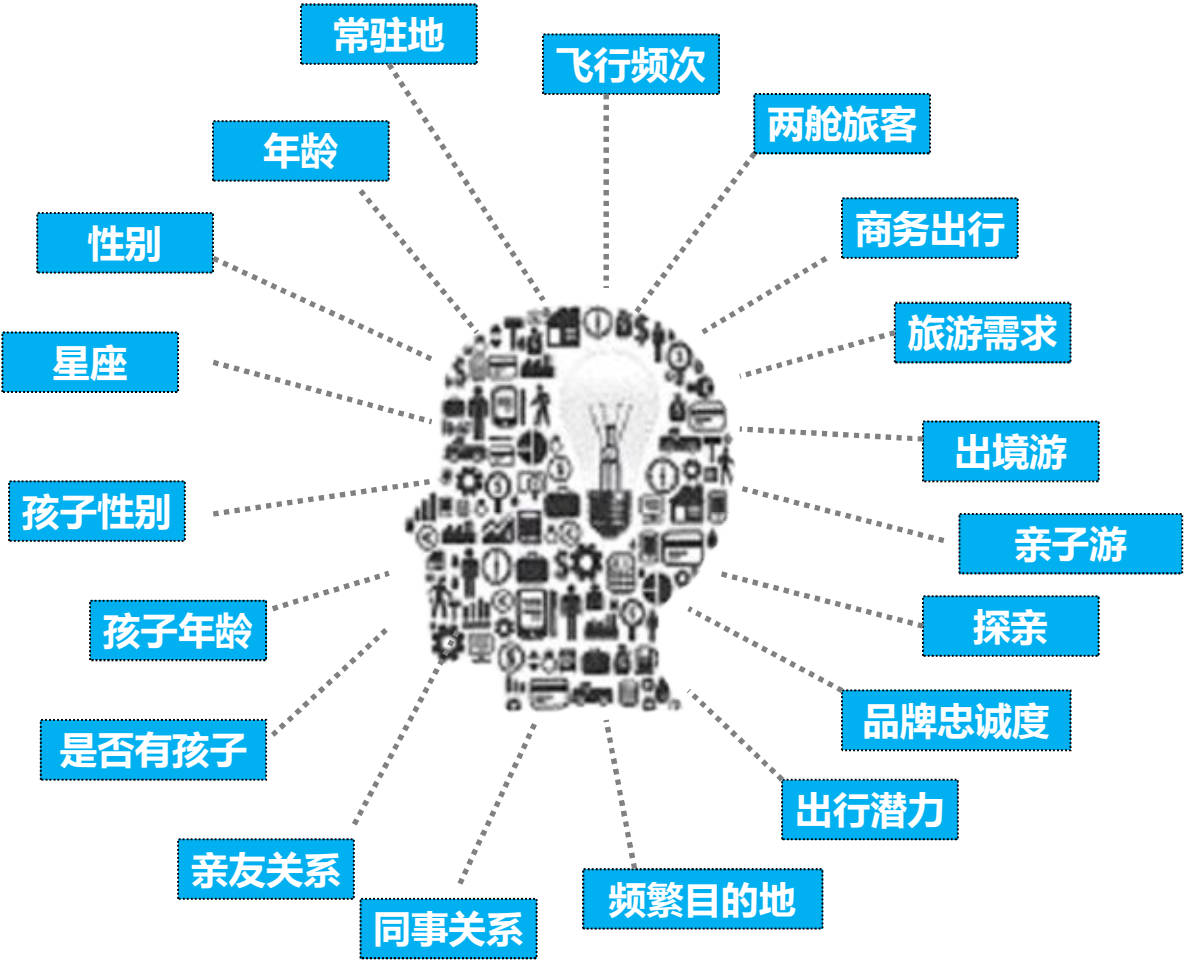
旅客数据统一视图



基础信息	客户信用	潜在需求
偏好	统一id客户 Profile	客户价值
同行关系	个体标签	群体标签

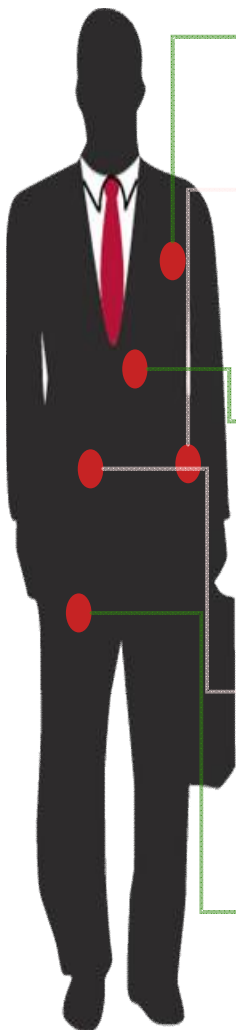


民航旅客画像





民航旅客画像标签



王明；男；出生于1980-01；【国籍】中国；【户籍】中国北京；【常驻地】北京；身份证110101198101010534；国航VIP；【常客卡】CA10000000；【大客户号】CA-C2345；【手机】15210000000-北京，13811111111-上海

“旅游”的出行特征：【出行次数】4次；【出行时间】春节2次，国庆2次；【出行日期规律指数】0.8999；【提前订票天数】提前7~15天4次；【航空公司】CA 3次，HU 1次；【随行人员】亲戚1次，独自3次；【订票渠道】BSP2次，B2C2次；【目的地】欧洲2次，东南亚1次，海南1次；【折扣率】0.5~0.8区间2次，0.8~1.0区间2次；.....

“北京-上海”航线：【出行次数】10次；【出行时间】平时8次，周末2次；【出行日期规律指数】0.6543；【提前订票天数】提前0~7天10次；【航空公司】CA7次，MU3次；【随行人员】独自1次，同事9次；【订票渠道】BSP8次，B2C2次；【折扣率】0.5~0.8区间2次，0.8~1.0区间2次；【舱位】F舱1次，C舱2次，Y舱7次；.....

“北京-广州”航线：【出行次数】3次；【出行时间】平时2次，周末1次；【出行日期规律指数】0.3213；【提前订票天数】提前0~7天3次；【航空公司】CA1次，CZ2次；【随行人员】独自1次，同事2次；【订票渠道】B2B1次，B2C2次；【折扣率】0.8~1.0区间2次，1.0~1.3区间1次；【舱位】F舱1次，C舱1次，Y舱1次；.....

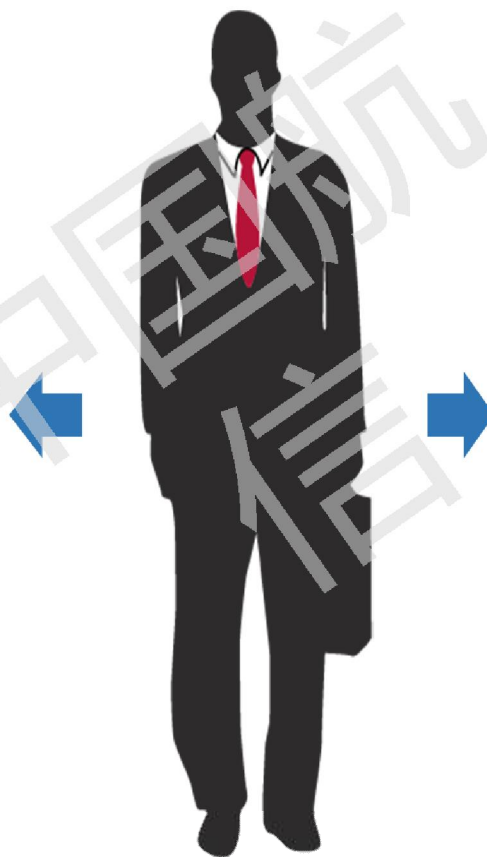
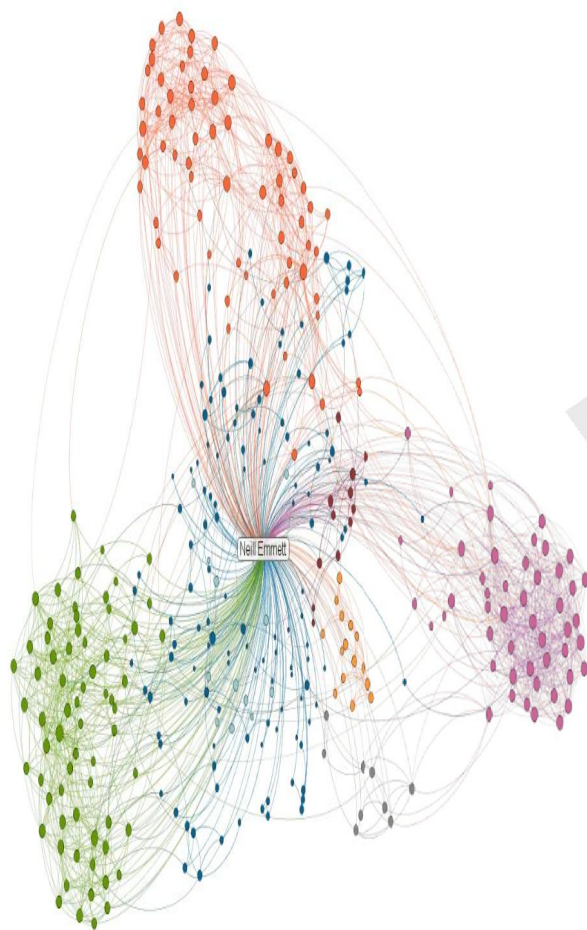
“春节”出行特征：【出行次数】2次；【出行目的】探亲2次；【随行人员】亲戚2次；【目的地分布】福建1次，江西1次；【目的地地停留时间】0~7天1次，7~14天1次；

“国庆”出行特征：【出行次数】2次；【出行目的】旅游2次；【随行人员】独自2次；【目的地分布】云南1次，四川1次；【目的地地停留时间】0~7天2次；

【关联目的地预测】欧洲：20%，东南亚：10%，美国10%；【国内飞行航段数】20次；【国内有效飞行航段数】19次；【国内平均折扣率】0.79；【国内平均消费金额】1345；【国内平均飞行里程】1329；【国际飞行航段数】2次；【国际有效飞行次数】2次；【国际平均客公里收益】0.16；【国际平均消费金额】2390；【国际平均飞行里程】3098；【第一次飞行时间】2010-03-25；【最后一次飞行时间】2014-07-30



旅客画像-全方位认识旅客



王硕

一共飞行了 **63239** 公里

77
综合价值

他的故事

50 次旅行

11 次与人同行

6 次大团出行

♂ 男性

📅 80后 狮子座

📍 住在北京

已婚

靠窗

商务人士

提前6天订票

经济舱

最常目的地：上海

亲子游：1次/年

国航

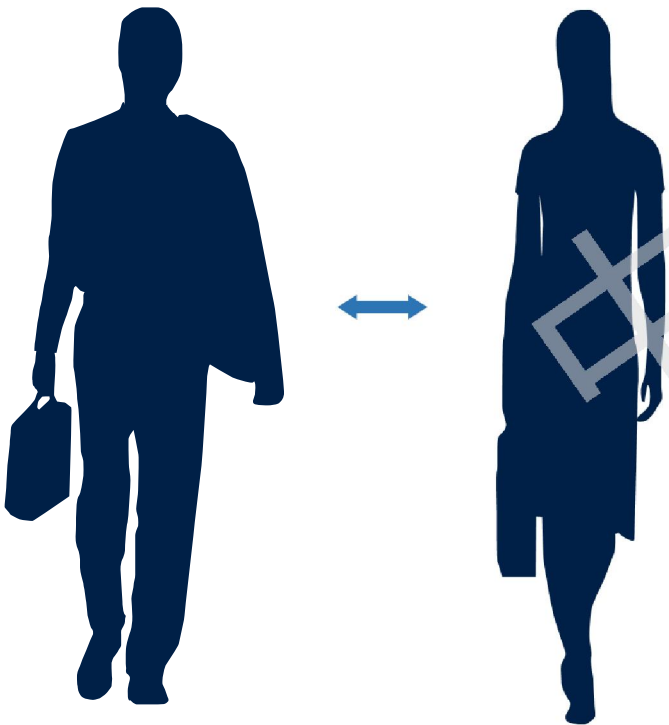
旅行：10次/年

官网购票

中国航信



旅客画像-同行关系发现



李雷和韩梅梅

一起旅行了 73 天

92 亲密度

他们的故事

12 次一起旅行

18 个共同好友

2 个共同群

♂♀ 他们是异性

👤 他们都是75后

📍 他们都住在北京

旅游

自由行

亲友关系

提前2个月订票

喜欢坐飞机后部并且靠窗

出境游

一起旅行：3次/年

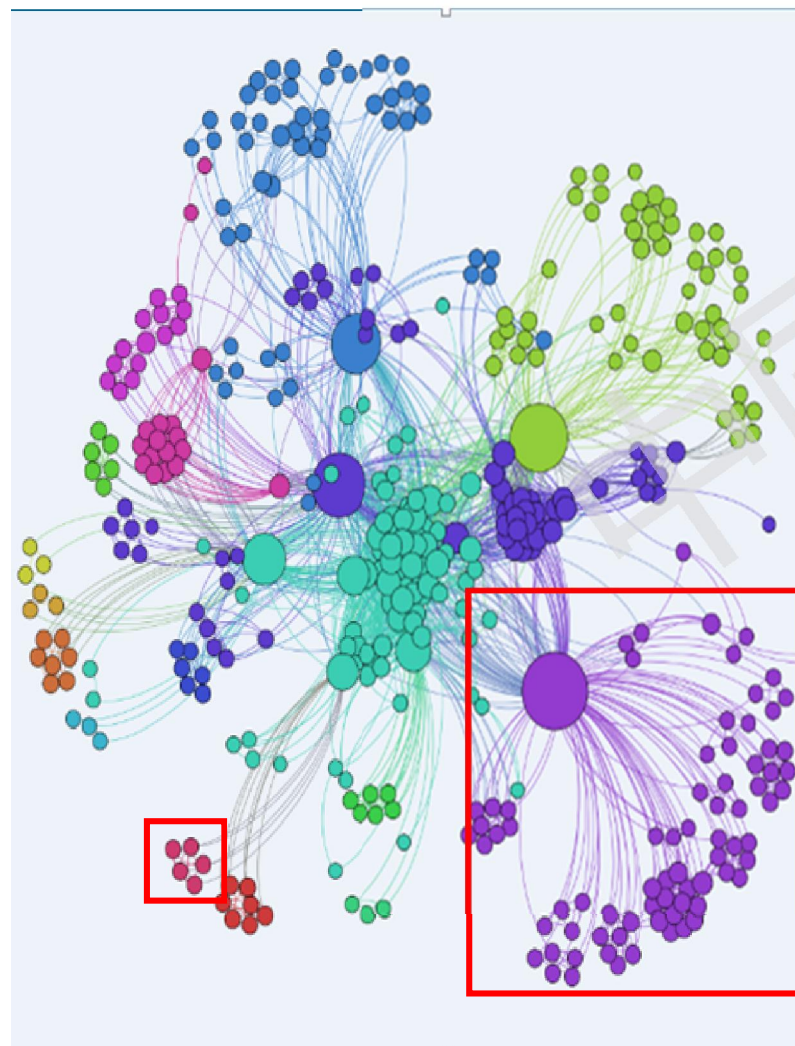
节假日

国航

手机值机



旅客画像-发现群体



Bruker中国

83
群体价值

年消费能力 **245** (万¥)

这个群的故事

23 位旅客

143 对同行旅客

3 次大团出行

♂♀ 男女比例8:2

👤 80后为主

📍 住在上海

商务群

无固定出票代理

金卡及以上旅客：8位

平均折扣：8折以上

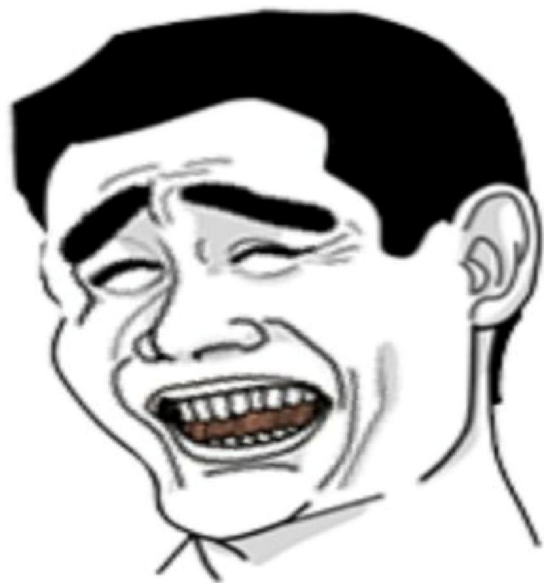
东航

旅行：420人次/年

最常目的地：北京

提前3天订票

经济舱



AIR CHINA
中国国际航空公司

您的位置: [首页](#) > [预订行程](#) > [航班选择](#)

北京 -- 洛杉矶 2017年10月20日 星期五 (97个航班, 3个直飞) 🔍 [更改搜索](#)

筛选条件	中转方式: ☐ 直飞 ☐ 1程中转 ☐ 2程及以上	起飞时段: ☐ 上午 (6-12点) ☐ 下午 (12-18点) ☐ 晚上 (18-24点)
航线特点	热门中转点 行程时间短 + 多选	
特色服务	机上娱乐措施 可充电 超大腿部空间 有wifi + 多选	

姚明 先生您好, 为您个性化推荐以下航班:

推荐指数	航班	出发	到达	飞行方式	特色服务	票价
★★★★	CA887 波音787 (大)	12:00 → 首都国际机场 T3	09:00 洛杉矶国际机场 B	① 12h 直飞		3289起 预定
★★★☆☆	CA987 波音787 (大)	15:00 → 首都国际机场 T3	12:00 洛杉矶国际机场 B	① 12h 直飞		3349起 预定
★★★☆☆	CA521 波音777 (大)	21:00 → 首都国际机场 T3	18:00 洛杉矶国际机场 B	① 12h 直飞		3435起 预定

其他搜索结果:

CA123--CA543 2程航班	11:20 → 首都国际机场 T3	12:20 洛杉矶国际机场 T3	① 16h 经停西雅图		2715起 预定
----------------------	--	---	-------------	--	--

AIR CHINA
中国国际航空公司

您的位置：首页 > 预订行程 > 航班选择 > 机票预订

航班详情

10-20 周五 北京 — 洛杉矶

中国国航CA887 波音787(大) 经济舱

北京时间
12:00
首都国际机场T3

12h
→

当地时间
09:00
洛杉矶国际机场B

[✈️ 退改签](#) [🧳 行李额](#)

购买建议：

等一等

未来走势：

我们预测本航线从今天起到起飞前，未来票价可能会更低，建议您等待时机再买哦！

舱位选择

头等/公务舱	超级经济舱	经济舱
☒ ￥21049	☐ ￥12429	☐ ￥3289

为您推荐

☒ 付费选座
优质座位全掌控
 ￥50

☒ 贵宾休息室
雅致高端 尊贵体验
 ￥60

☒ 预付费用行李
行李费用预付
 ￥100

[更多附加服务](#)

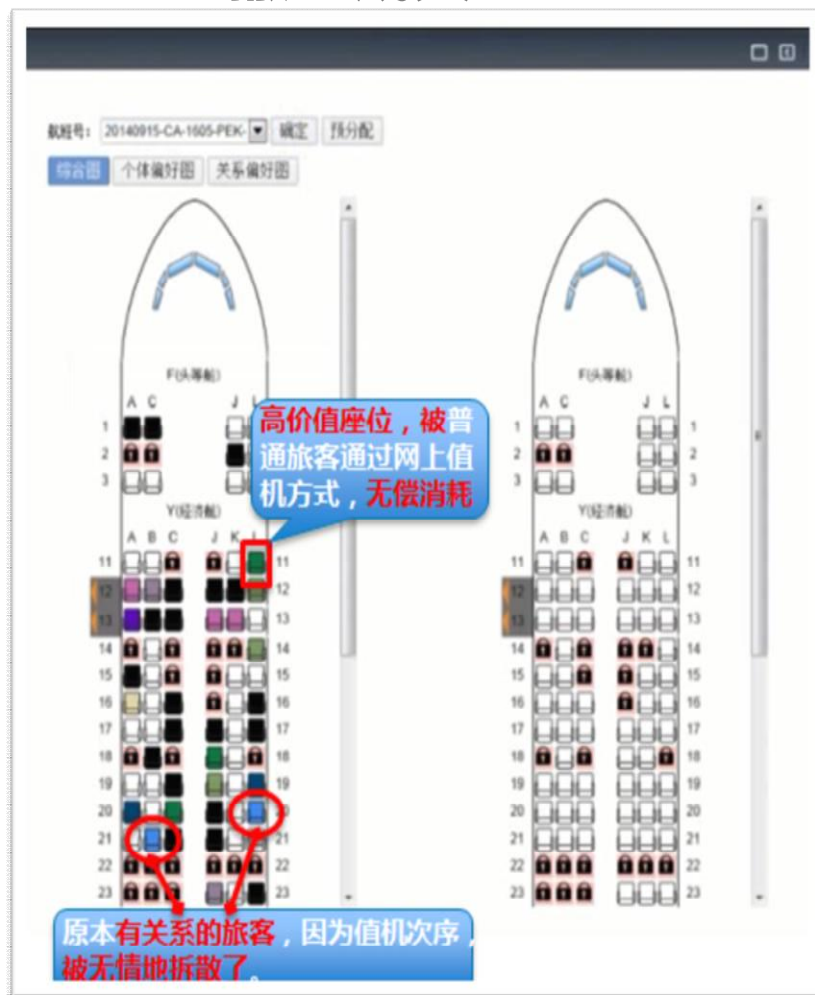
[下一步](#)



旅客画像应用2-智能座位推荐

过去

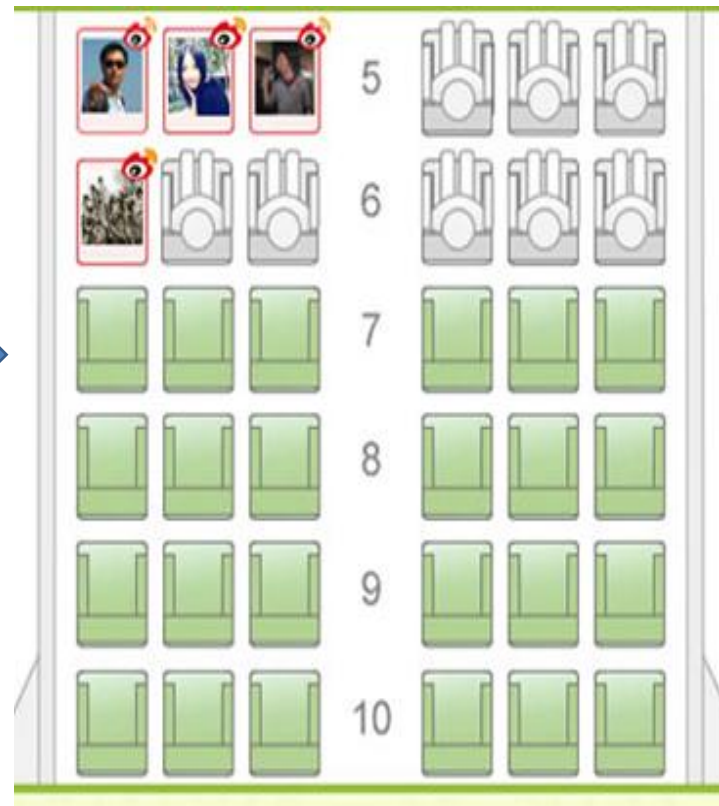
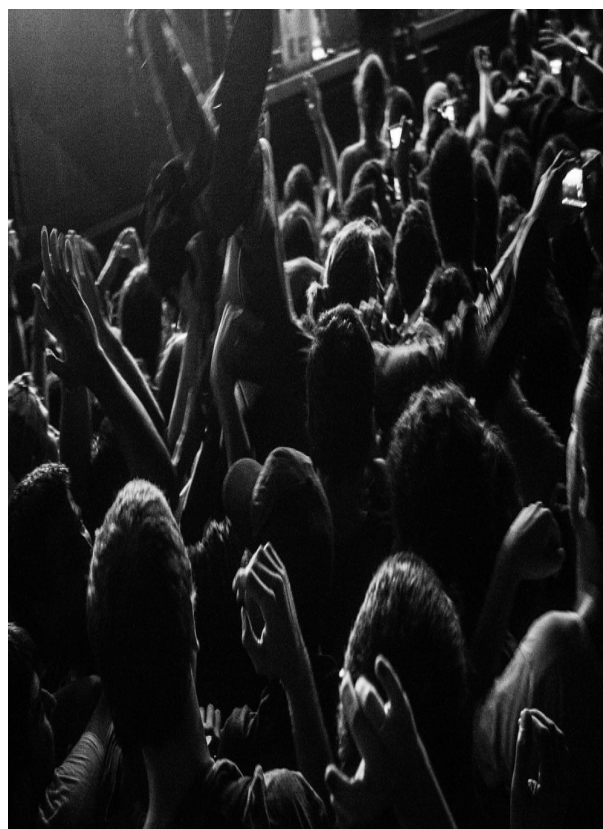
撒豆成兵



现在

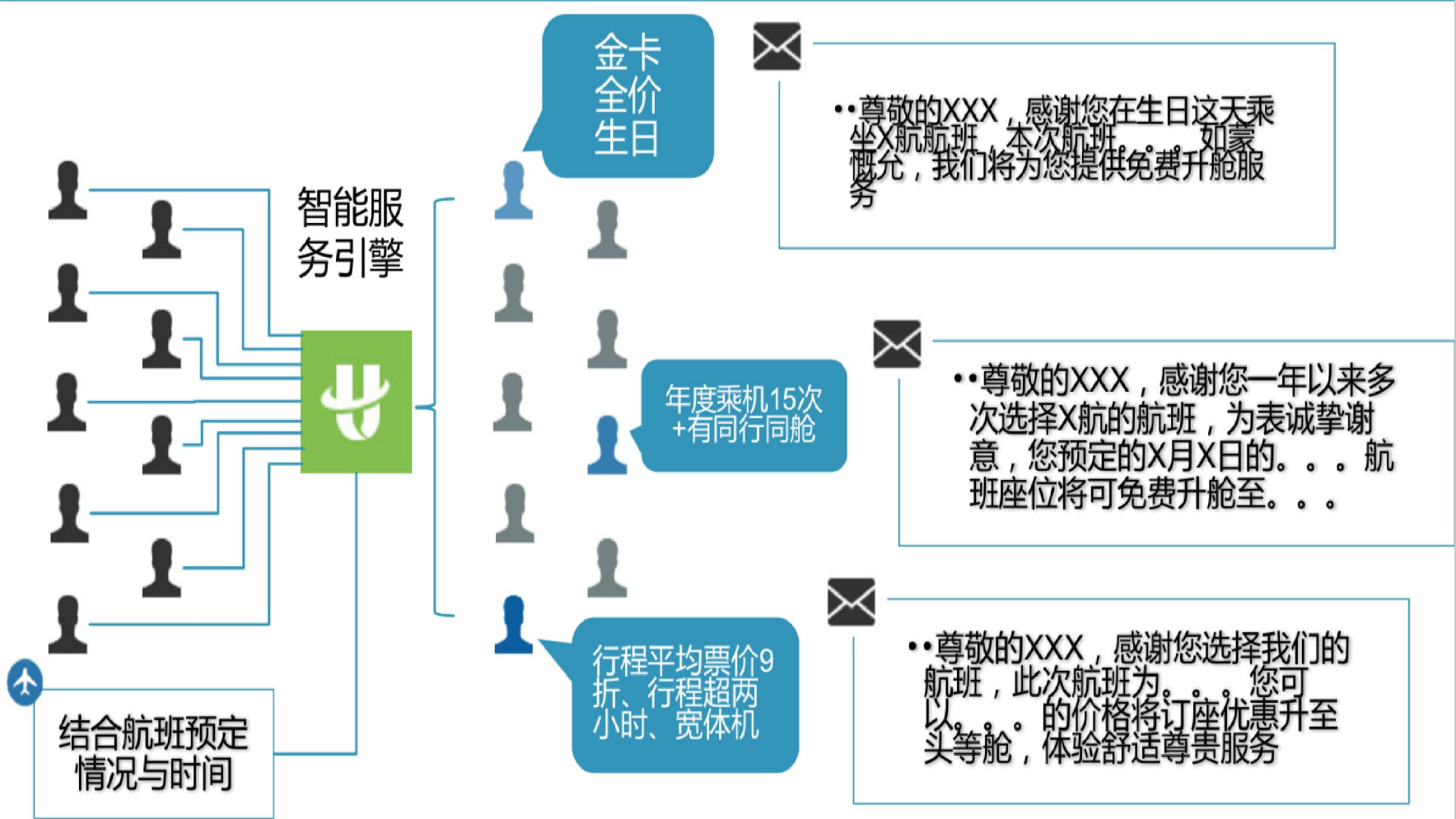
旅客认知：个体偏好/价值+群体关系





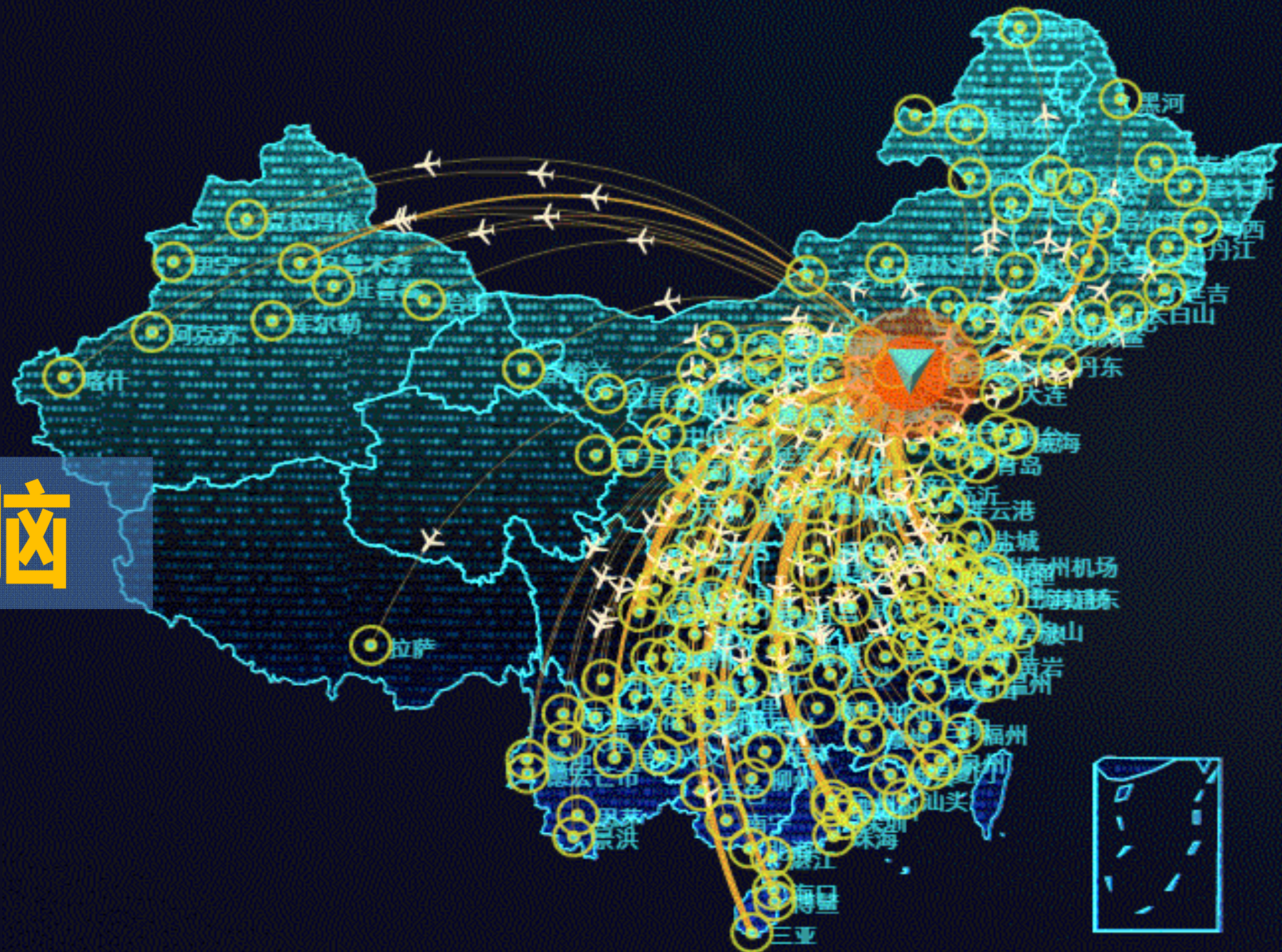
方式A：航旅纵横按航空公司所制定规则筛选出目标旅客，向其定向推送航空公司优惠或服务产品

方式B：航空公司给出目标旅客名单，航旅纵横向其定向推送航空公司优惠或服务产品



升舱、付费座位、选餐、机场接送、泊车等

2.民航大脑





- 建立指数分析模型，以指数化、可视化的方式解读行业数据，面向政府机构和行业专家
- 致力于为民航业、旅游业等相关行业和政府机构提供最权威、最及时、最具有价值的数据统计与分析成果
- 运用3D数据可视化、人机交互等技术，多角度深层次的阐释数据应用的价值。



[2018 春运分析报告]

2018 年春运期间，民航市场保持平稳有序的发展。今年春运与去年相比，市场整体运行情况良好。国内、港澳台和国际市场出行旅客均创新高，体现了民航市场的强劲生机和发展动力。由于整体市场运力投放增幅高于旅客量增幅，各市场客座率均保持平稳或小幅下降。除综合市场外，平均价格略有下降。今年春运期间的旅客流向仍呈现由南向北的特点，节前返乡、节后复工，但由于春节时间较晚，节后旅客量与节前相比，集中形成错峰高峰，给民航运输带来了不小的挑战。重点航线、机场旅客量均有所上涨，增幅各不相同，形成了不同的发展态势。]

【中国民航信息网络股份有限公司航指数】
【北京市顺义区后沙峪中国民航高科技产业园区】

2018 年 11 月 15 日



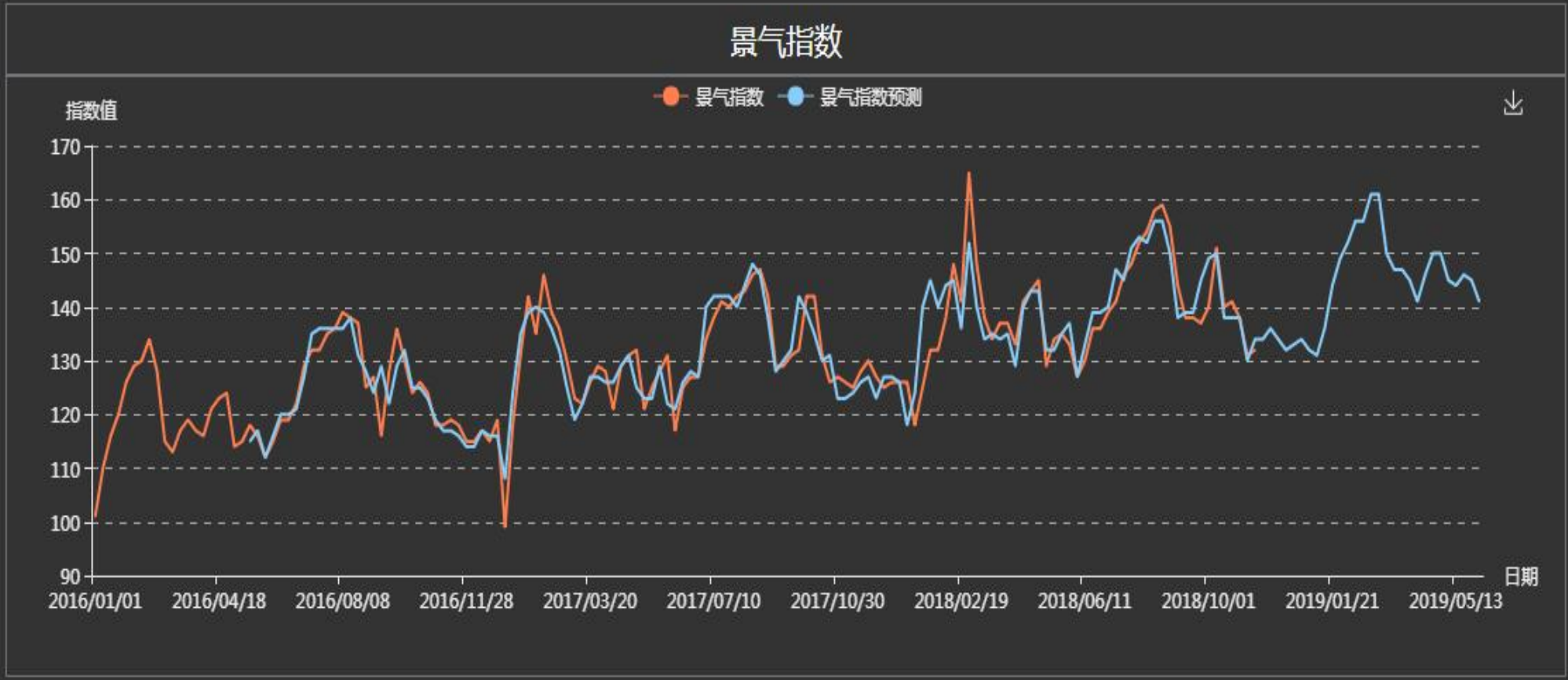


中国民航景气指数分析预测

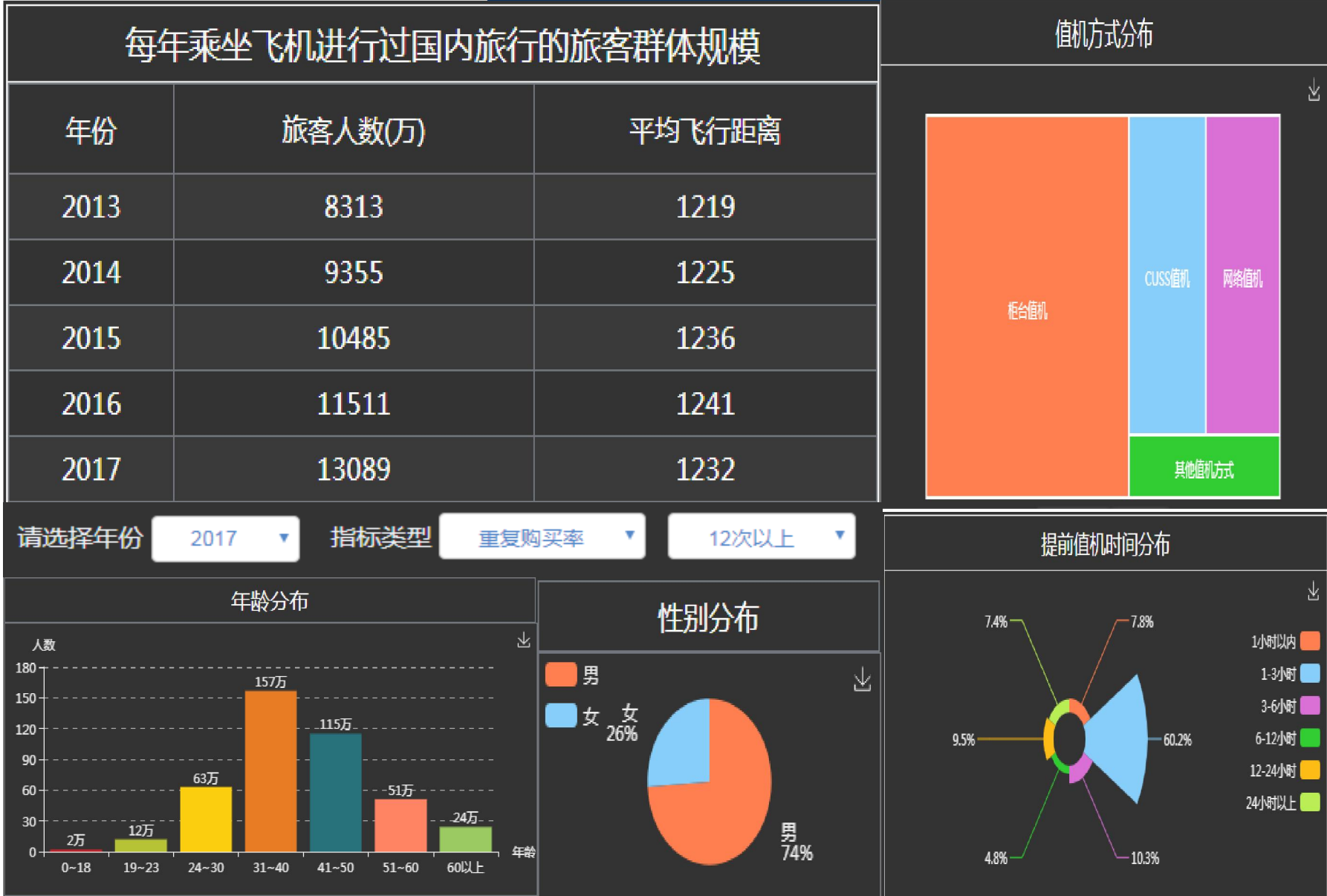
景气预测 ?

请选择航线类型：国内 国际 港澳台

景气指数

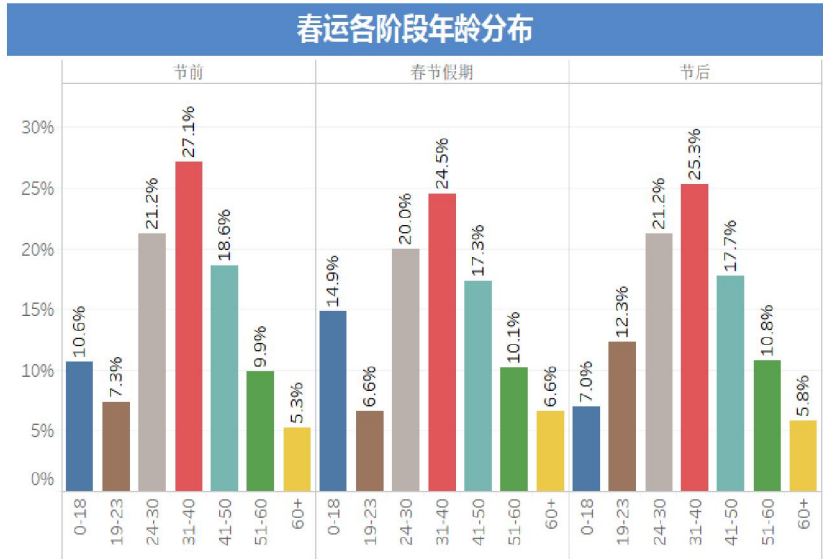
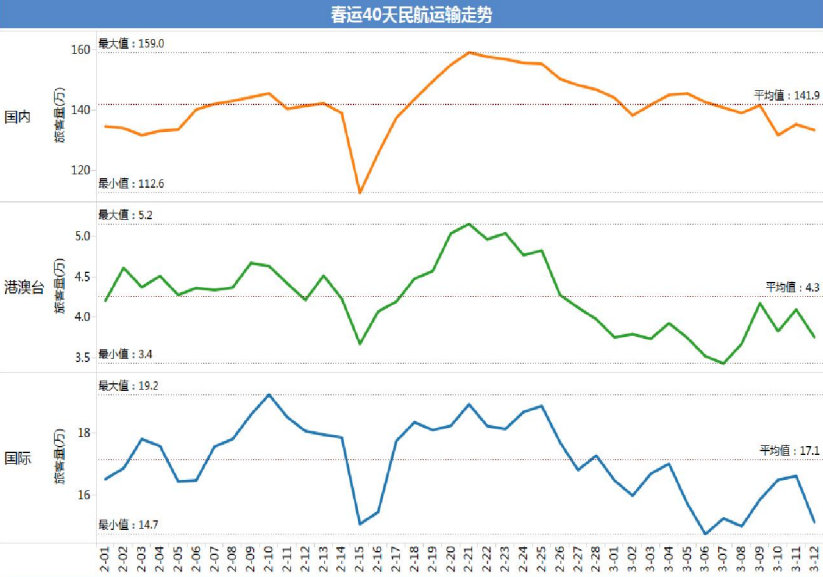


中国民航市场旅客特征分析

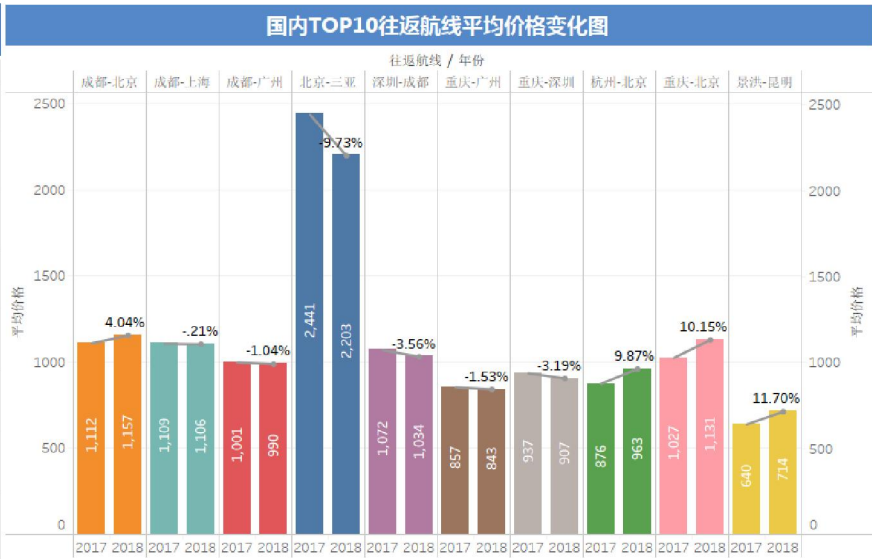
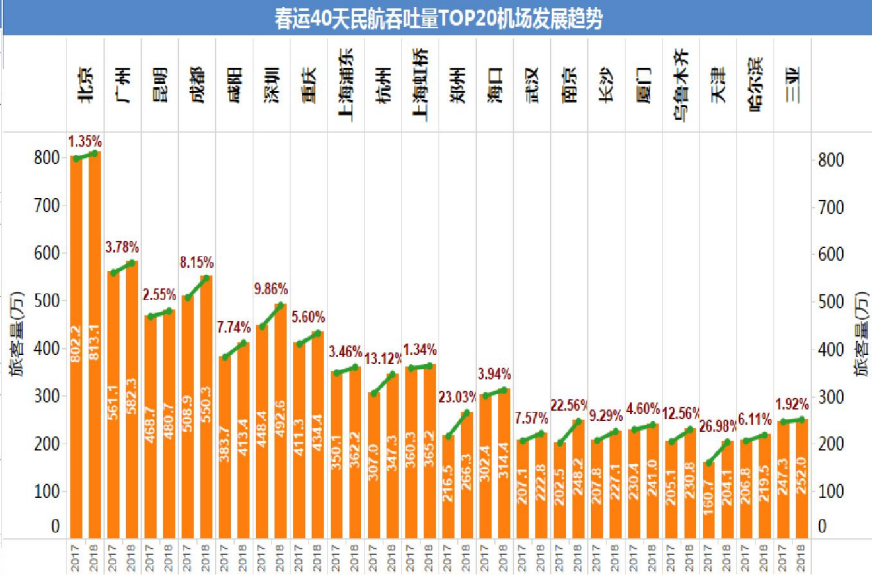




2018年春运四十天分析



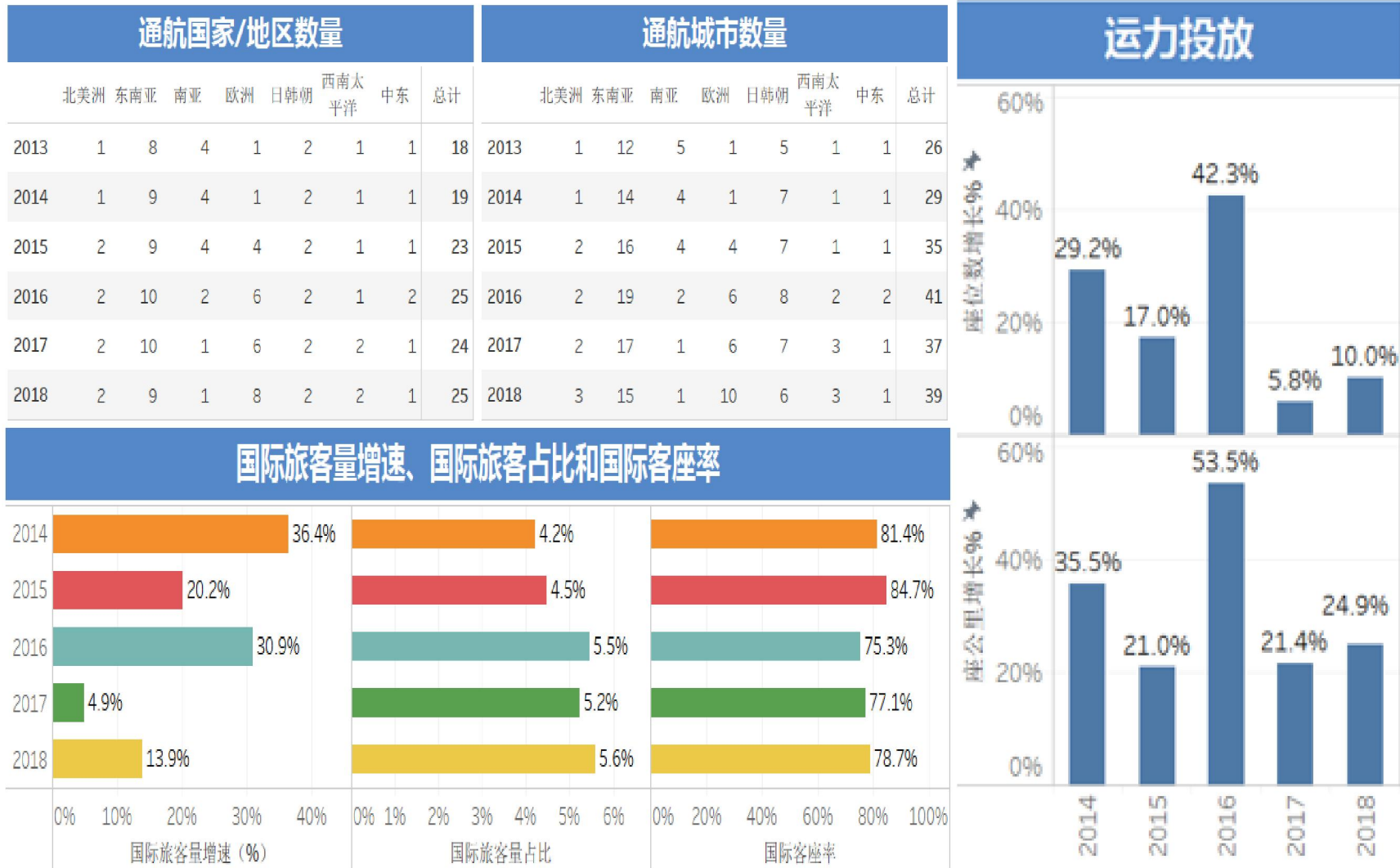
- 春运航线旅客量增幅整体呈现“前低后高”的状态
- 80后和90后占据民航市场半壁江山，春节期间商务人群24-50岁各阶段人群占比下降明显



- 重点机场吞吐量中西部枢纽机场发展快于东部枢纽机场
- 不同航线的平均价格走势各不相同，部分潮汐属性航线平均价格降低，大部分旅游属性航线平均

国际枢纽建设观察——

成都国际民航市场蓄势待发



- 积极布局洲际航线，欧洲航线扩张明显
 - 国际旅客连年增，客座率水平渐改善
- 座公里增长迅速
 - 中远程航线快速扩大



目录



1

民航进入大数据时代

2

民航旅客服务智能化实验室

3

实验室应用

4

互动环节(Q&A)

感谢大家！

