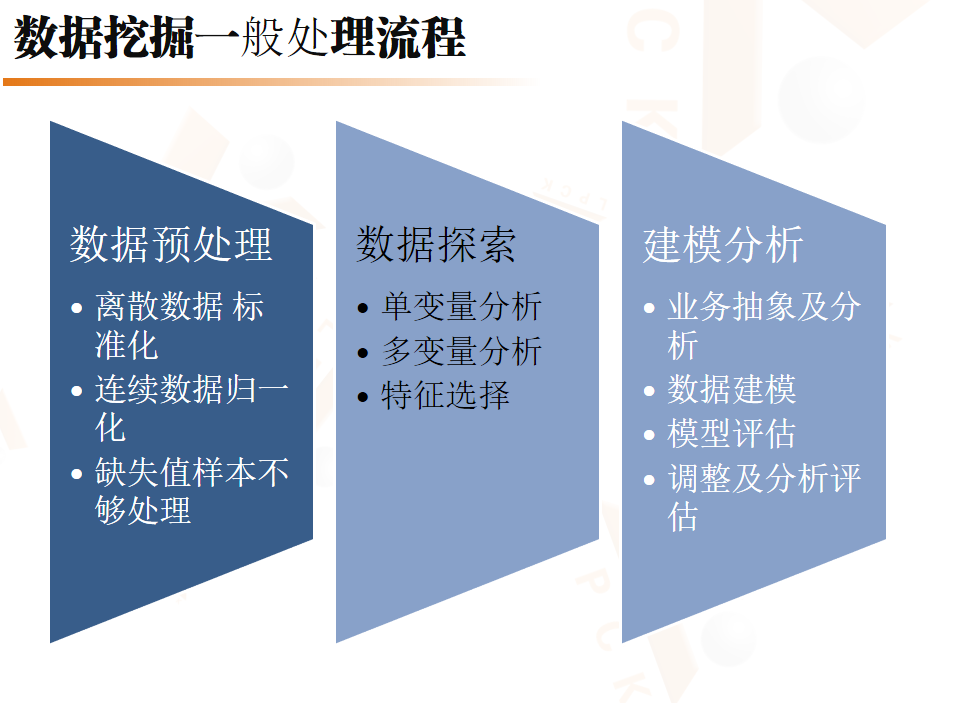
## 3 数据分析挖掘

根据数据挖掘一般处理流程， 完成数据清洗之后,接下来对数据进行探索.



### 3.1单维分析

具有海量千万级汽车行业车型、配件数据库和国内领先的汽配价格处理系统，可提供更加强大的数据服务支持，结合保险公司现有数据库，可以对各个品牌、品质汽车配件进行通过上全面的分析

#### 3.1.1空间维度分析

空间维度分析:根据市级公司、省级公司、区域中心、全国4个维度对配件价格进行分析。

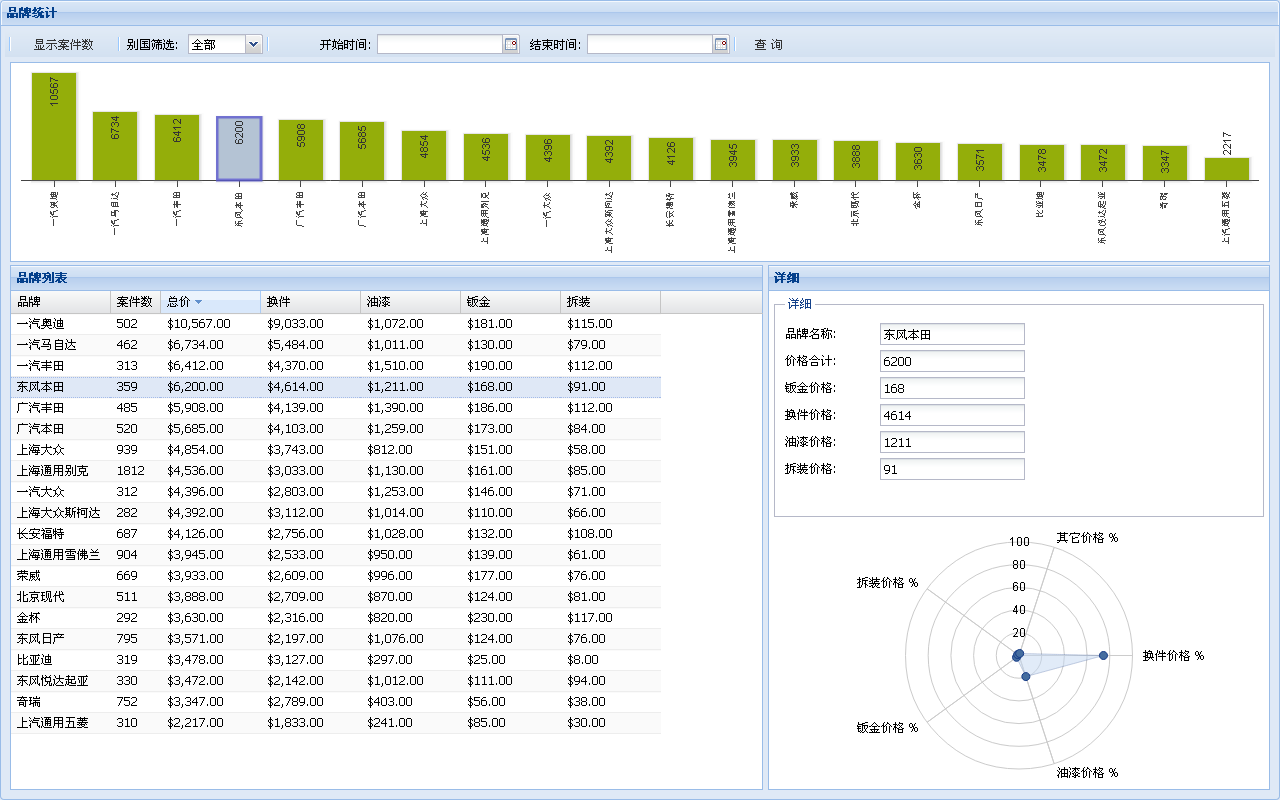


#### 3.1.2时间维度分析

在对应空间维度的基础上，提供实时、月度、季度、年度和长时间维度的配件价格分析。

#### 3.1.3.车型品牌维度

对于不同品牌的配件进行配件价格的分析。进而分析相对风险指标、辅统计各品牌出险次数、案均总价、案均换件维修费用及各部分占比。助宏观理赔管理。



### 3.2多变量分析

在单维分析的基础上，对空间、时间、车型品牌等多维度、动态地分析配件价格波动和品牌更新情况，形成清晰简明、合理实用的分析报表和统计分析图形：其他维度：根据具体单个修理厂、供应商或者不同配件品质进行价格分析。

### 3.3假设分析

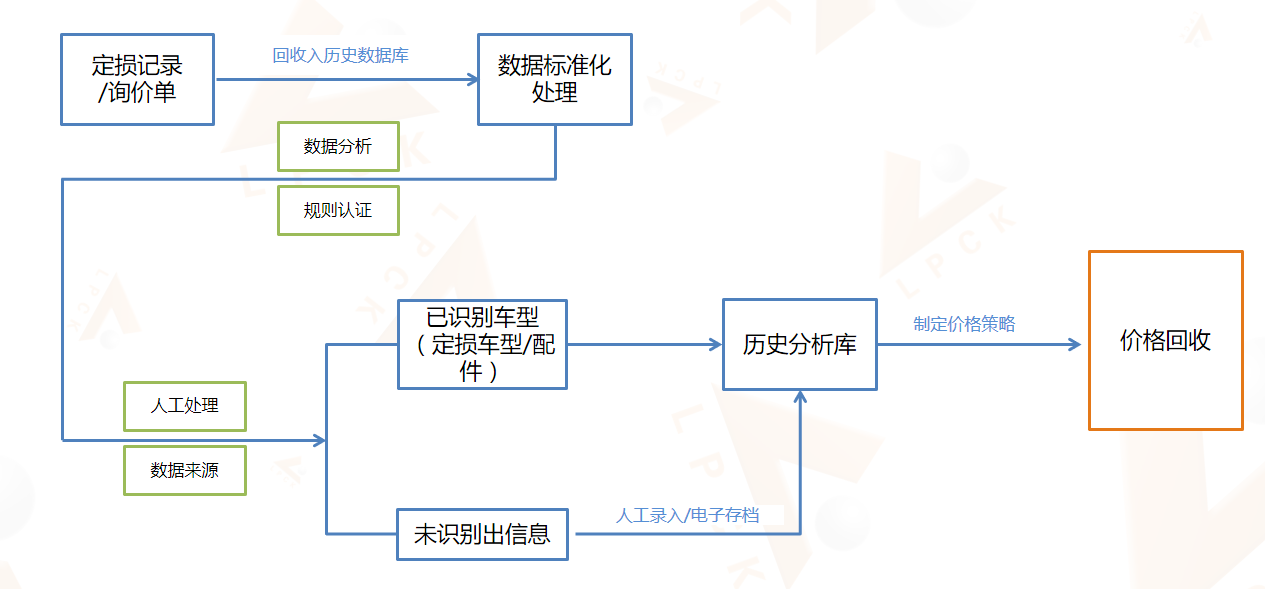
在单变量和多变量分析基础上，可以设定时间、空间、车型品牌、维修厂等条件，并且允许用户设定符合实际需求假定条件，如配件价格的平均值、最小/最大值，中位值、Q1/Q3、区间、价格分布类型分析、易损件目录分析，观察条件成立的情况下相关分析或指标的变化，最终按照客户实际需求输出相应的即席查询报表、动态分析报表，使得在对接定核损工具和询价供货平台时，对业务实际操作具有实时指导意义。

### 3.4 建模分析

在多维度数据分析基础之上，在空间、时间、车型等不同的维度，运用数学统计模型建立相应配件提出实时配件指导价模型，提出配件指导价格，从而指导和规范分公司和报核价区域中心的报核价操作，管理和控制业务风险水平。

#### 3.4.1价格回收

根据已有定损数据和询价数据，该数据存入配件价格历史数据库，进过标准化处理，对标已有的车型进行自动处理；如通过进入历史分析库。如不通过人工进行修正或进入电子存档。对进入历史分析库的数据进行分析，根据制定的价格策略规则成为最新配件价格影响因子。



#### 3.4.2价格模型

1.当数据样本量充分时：

通过对历史数据的大数据分析，形成多维度分析，并运用时间序列分析等数学统计模型对配件价格建立数学模型，具体步骤如下：

①用统计方法取得配件价格的时间序列动态数据。

②根据动态数据作相关图，进行相关分析，求自相关函数。相关图能显示出变化的趋势和周期，并能发现跳点和拐点。跳点是指与其他数据不一致的观测值。如果跳点是正确的观测值,在建模时应考虑进去,如果是反常现象，则应把跳点调整到期望值。拐点则是指时间序列从上升趋势突然变为下降趋势的点。如果存在拐点，则在建模时必须用不同的模型去分段拟合该时间序列，例如采用门限回归模型。

③辨识合适的随机模型,进行曲线拟合,即用通用随机模型去拟合时间序列的观测数据。对于短的或简单的时间序列，可用趋势模型和季节模型加上误差来进行拟合。对于平稳时间序列，可用通用ARMA模型（自回归滑动平均模型）及其特殊情况的自回归模型、滑动平均模型或组合-ARMA模型等来进行拟合。

运用时间序列模型进行线性最优预测与控制，从而制定合理、严谨的推荐指导价格。

1. 当数据样本量不足时：

可以通过放宽时间、空间、车型等维度扩充数据，再了、利用时间序列分析模型进行价格预测。

3.当样本数据稀少且价格波动较为平稳时：

提供多种价格参数估计，包括点估计、区间估计。

4.当样本数据稀少且价格波动幅度较大时，可直接进行人工干预。

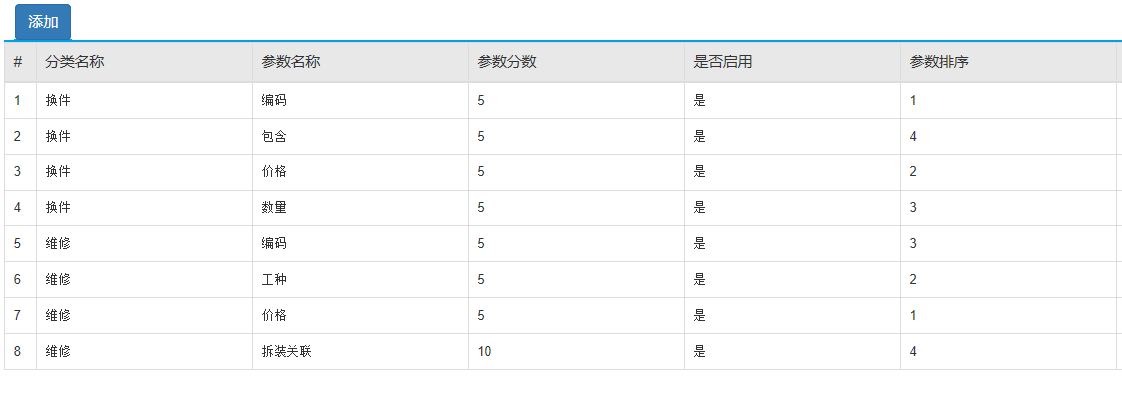
减损分析：

如果参照配件价格模型指导核损，可以有效减少定损金额，相对于系统配件价格，可以减少

pi为

#### 3.4.3风险**模型**

利用价格模型，对于超出范围的异常情况进行风险管控和警示，从而指导和规范分公司和报核价区域中心的报核价操作，管理和控制业务风险水平。 通过大数据平台的学习，挖掘定损单中配件和配件同时出现的比例，配件和配件内在关联性，配件和工时关系，配件和车型的关联性，配件的价格合理性等，并且计算规则和规则之间的权重，最终对于定损单中相关的风险进行评估和告警。





#### 3.4.4整案模型

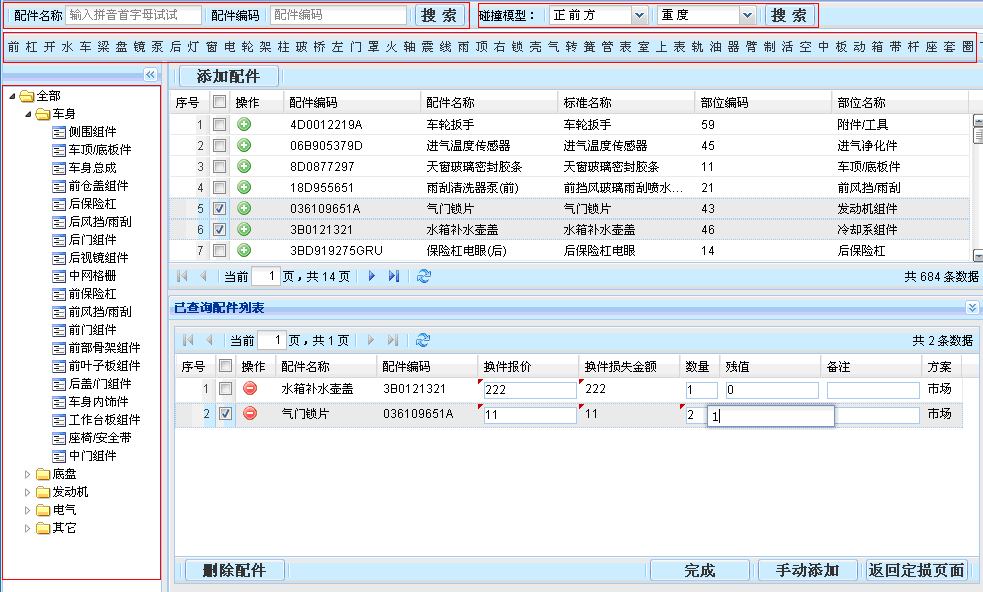
在保险行业车险理赔相关领域，一直都存在定损难、管控难等问题，信息标准化尚未在车险行业普遍应用是其主要原因之一。对于理赔定损数据的分析，仅从提取案件特征、多维度分析并不能十分准确地反映整个案件所呈现的全部信息，更难以将各个案件之间的隐含关系联系起来。因此，从整案分析的角度出发，对于案件的分析、数据的分析都有着极大的指导作用。具体可分为以下四个方面：

1. .对理赔作业过程中的口语化不规则数据，进行标准化处理，提取事故损失特征向量，通过历史样本的挖掘分析及理赔信息的内部约束关联等进行逻辑推理，建立一套多级关联的车险事故损失模型，并为每个模型给出提供合理的理赔金额区间。基于此套处理体系建立的智能化平台，可为车险理赔行业的定损、审核等环节提供更加精确高效的后台作业支撑。

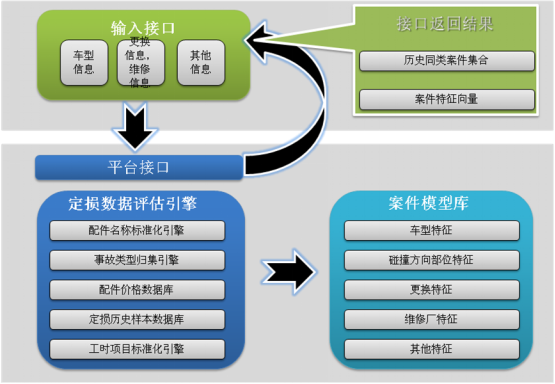


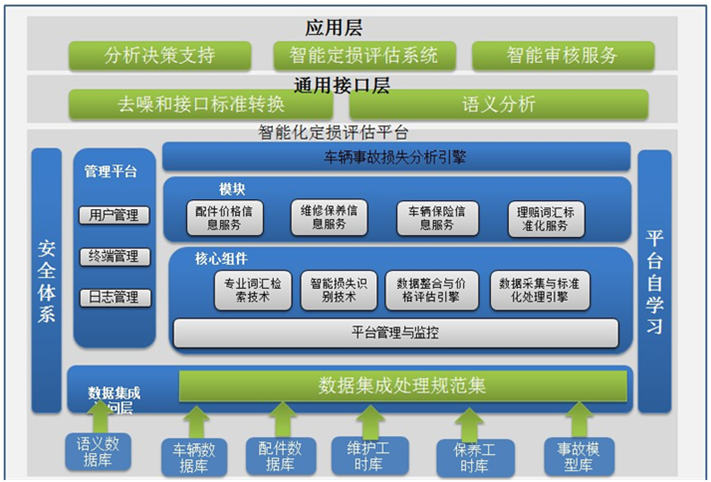
1. .在事故信息标准化方面，通过对千万级的车险、汽配、汽修相关专业词汇进行充分的分析归纳，提出了一个基于语义网络、跨领域的关联逻辑推理方法，通过引入本体特征字符集，构造输入输出之间的映射空间，使原始输入的不规则事故信息在语义层面上得到精准的智能化匹配描述。比如输入“qb”（前保）显

示的结果是前保险杠；输入“yj”（羊角）显示的结果是转向节。



1. .在事故损失模式识别方面，对事故损失样本提取出一套符合车险事故损失情况的特征向量集。对历史样本进行特征选择和迭代学习，并应用贝叶斯网络分类器进行类别输出，从而匹配出历史同类型案例，用以支持理赔结果的风险识别及控制。





#### 3.4.5整案分析

赔估值方面，根据系统内置的定损运算模板结合专家经验值，对海量实际数据进行监督训练，改造出适用于车险理赔估值的神经网络算法，并通过敏感性分析和自动校正，得到符合实际的事故理赔估值置信区间,用以支持定损作业及审核决策。



### 3.5建立动态指标

结合配件价格多维分析结果、指导价格模型，建立分公司各项动态指标：汽车配件价格指数、配件价格异常比例、配件引用频率、、配件车价比指数等指导性指标，观察监控理赔配件价格的变化和相关分公司政策执行情况，形成KPI指标展示各分公司配件理赔优劣情况。

#### 3.5.1汽车配件价格指数

汽车配件价格指数是参考了股票价格指数的编制原理，是反映不同时期一组商品（服务项目）价格水平的变化方向、趋势和程度的经济指标，通常以报告期和基期相对比的相对数来表示。具体步骤如下

维度表建立**：**提取两年的年度数据作为基准表，选取常出险前10个品牌车和前20种配件

数据刷选：剔除一些异常值（价格异常高或异常低），保留平均价格上下30%的数据：

配件价格确定：对同一月的同品牌的同一种配件的多次价格的记录求均值作为该品牌该配件在这个月的价格，由此可以得到不同品牌所有配件在这个月的价格。

配件权重的确定：以不同配件换件次数的占比（即不同配件的换件次数/总换件次数）作为不同配件的权数。

配件价格指数算法：



*P1*为报告期所有抽选的n种商品的价格,*p0*为基期这些商品的价格，q为权重。

#### 3.5.2价格异常比例率

根据数据分析和价格模型，计算分公司区间内的配件异常数量，根据该数量计算异常配件在全部配件中的比例



nu表示异常配件数量，n 表示区间内总体配件数量。

#### 3.5.3价格异常总价比例率

由数据分析和价格模型，得出分公司区间内的配件异常比例及配件价格，计算出异常配件总价在全部配件中的比例



nu表示异常配件数量，pu 表示异常配件价格，n 表示区间内总体配件数量，p表示配件价格。

#### 3.5.4换件占比

计算配件换件费用占定损金额总费用的百分比；



P1表示配件换件费用，p0 表示定损金额总费用。

#### 3.5.5件均

在一定的时间段和定损价格区间内，平均每个案件的定损价格；（包括各保险公司的件均；整个行业的件均；维修厂的件均;定损人员的件均）（剔除低于100的，计算行业时，数据选取均值上下的30%）。

### 3.6挖掘隐含业务规律

利用多种数学的分析方法和数据挖掘技术建立模型，找出配件价格变动趋势及区域对比、异常区域识别等隐含的业务规律。为运营状况分析和决策支持提供洞见分析功能。

#### 3.6.1异常区域识别

通过对大量数据采用机器学习方法，如SVM、Random Forest，智能识别配件异常的维修厂、4S店或分公司。

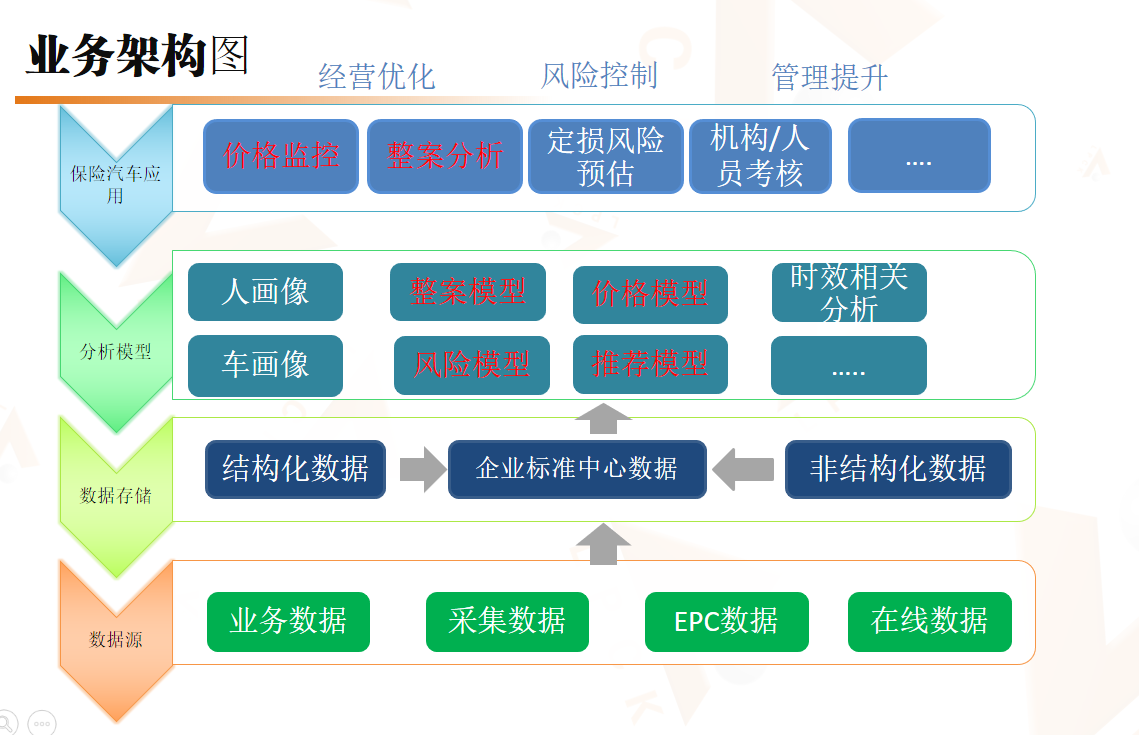
#### 3.6.2价格差异因素识别

运用多因素方差分析法，分析机构之间价格存在着显著的价格差异的主要因素。

#### 3.6.3. 出险配件排名

根据行业整体的数据，抽取数据排名，生成总公司的出险配件排名及各个保险公司的出险配件排。

### 3.7小结：



通过对历史数据的单维和多维度假设分析和挖掘，实现实时、多维、动态的汽车配件价格分析，生成在一定假设下的基于统计数学模型的配件指导价，指导和规范分公司和报核价区域中心的报核价操作，减少客户风险水平。通过数据挖掘技术挖掘隐含业务规律，观察监控理赔配件价格的变化，指导机构和理赔人员规范合理的执行配件价格理赔政策，为运营状况分析和决策支持提供洞见分析功能。

**第五章 其他信息**

**第六章 供应商情况介绍**

**1、公司简介**

浙江联保网络技术有限公司成立于1998年，注册资金貳千万元整，公司下有IT部、数据部、业务部、客服部及行政部，现有员工百余人。

公司下属机构包括：浙江机动车辆保险理赔调研中心， 《中国车险理赔参考》，浙江新东方保险公估有限公司。 国内最早的汽车定损信息咨询系统服务机构，公司始终专注在汽车数据、汽车保险理赔数据、和相关的系统平台领域，致力于提供符合中国实际市场需求的方案和服务，并且不断创新汽车行业与保险行业发展所需的信息服务平台，尤其在基于移动互联和商业智能的汽车保险数据应用平台方面。

公司总部位于杭州，并在上海、宁波等城市设有分支机构，面向全国开展业务， 主要提供数据、软件平台、软硬件开发、培训及公估服务。公司拥有国内外资深的技术专家和业务管理精英组成的综合性团队，精通汽车行业数据、车险行业系统、移动互联和商业智能，力求给客户提供专业可信赖的数据和系统服务。成功案例有中保信、上海保协、浙江省交通厅、大地、永诚、紫金、永安、安联等总对总合作，及上千家组织机构，包括物价局和汽车服务公司等。

**2、公司实力**

浙江联保网络技术有限公司成立于1998年，始于发行《理赔参考》专刊和汽配价格查询系统，为人保及其他公司基层机构提供专业的汽配数据服务。2012年，公司追随大数据发展时代，开始向数据科技型企业方向转型。建立业务管理软件、数据深挖掘、移动互联产品开发项目，并积极拓展汽车后市场领域。

2002年与浙江大学合作，研发出理赔定损管理软件，并推出中国车险服务网[www.lpck.cn](http://www.lpck.cn)。

2007年底，完成数据标准化工作，推出全新的LPCK定损系统3.0版本。

2008年推出行业领先的MAS移动定损系统，并在2010年随3G的正式商用，发布3G版的MAS，提供国内第一个全流程手机定损功能（GPS定位调度、接收案件和抄单信息、移动查勘、移动定损、打印出单、），以及装载于手机的海量数据库结合。整体效用处于行业领先地位。

2011年在上海保险同业公会推出行业领先的定损数据智能标准化引擎，获得行业多方好评。

2012年推出3G手机定损系统MAS5，与中国保信合作建立理赔大数据标准化平台。

2013年公司正式发布业内领先的智能汽车定损评估引擎。

2014年推出维修工时数据系统，数据智能标准化匹配和评估引擎。

2015年和平安、太平合作启动，精细配件数据；车后市场合作启动。

2016年和中保信北分合作启动，建立理赔大数据标准化平台。

目前与超过600家保险省分或者中支公司保持数据合作，基层反映效果良好。

我司现有专业开发人员20人、数据分析挖掘人员15人、和业务数据专家20人，90%以上的业务都是和车险及车辆配件相关的项目，足以满足CPIC的需求。

1. **代表性项目**

|  |  |  |  |
| --- | --- | --- | --- |
| 项目名称 | 上海车型信息平台理赔信息服务系统零配件标准数据库项目第一期 （中保信） | 上海车型信息平台理赔信息服务系统零配件标准数据库项目第二期 （中保信） | 北京车险信息平台理赔车辆修理数据治理服务项目合同（中保信） |
| 所在行业 | 保险 | 保险 | 保险 |
| 案例的目标和范围 | 整个上海市 | 整个上海市 | 整个北京市 |
| 覆盖的组织范围 | 上海市所有保险公司 | 上海市所有保险公司 | 北京市所有保险公司 |
| 实施时间 | 2016.4 | 2016.5 | 2016.9 |
| 案例实施的结果 | 配合行业协会和保监局做了很多分析和指标，对上海保险行业的数据进行清理，标准化程度达到99.7% | 配合行业协会和保监局做了很多分析和指标，对上海保险行业的数据进行清理，标准化程度达到99.7% | 配合行业协会和保监局做了很多分析和指标，对北京保险行业的数据进行治理，效果显著 |
| 供应商在案例中所承担的职责 | 数据清理及配件标准化 | 数据清理及配件标准化 | 1平台搭建；2数据标准化处理；3数据抽取。 |

1. **团队成员**

项目经理：吕建丰

职责：对理赔配件监控平台具体事务内容的安排及部署。

系统架构师：卢伟舜

职责：系统构建。

开发工程师：徐君、代和强

职责：对理赔配件监控平台前端及后端系统开发。

数据分析挖掘人员：王斐成、王涛、王琉立

职责：主要负责对数据清洗，对数据进行抽样性数据分析，以及对数据的整理，从多个维度对数据进行抽样数据分析挖掘。

测试负责人：徐琮刚

职责：负责理赔配件监控平台的测试工作。

测试及其他支持：朱煌菲

职责：理赔监控平台的测试及其他支持。

业务数据专家：贾健、莫立利

职责：对业务数据进行定标指导，还有一定程度的人工干预。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 吕建丰 | 工作年限 | 10 | | 专业 | 计算机科学与技术 |
| 毕业学校 | 浙 江 大 学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 汽车配件二代系统、货运车辆保险数据模型及校验服务项目 | | | 项目经理 | |  |
| 二 | 中保信上海数据清理标准化 | | | 项目经理 | |  |
| 三 | 配件包含关系基础数据（平安） | | | 项目经理 | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 卢伟舜 | 工作年限 | 10 | | 专业 | 计算机 |
| 毕业学校 | 浙 江 大 学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 汽车配件二代系统 | | | 架构师 技术负责人 | |  |
| 二 | 保信数据清理标准化 | | | 技术负责人 | |  |
| 三 | 大地移动定损系统  永安定损理赔  杭州污水处理系统 | | | 架构师 | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 王斐成 | 工作年限 | 8 | | 专业 | 计算数学 |
| 毕业学校 | 国防科技大学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 汽车配件二代系统 | | | 数据分析师 | |  |
| 二 | 保信数据清洗及标准化 | | | 数据分析师 | |  |
| 三 | 货运车辆保险数据模型及校验服务 | | | 数据分析师 | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 王涛 | 工作年限 | 2 | | 专业 | 信息与计算科学 |
| 毕业学校 | 浙江财经大学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 北京行业数据清理 | | | 数据分析师 | |  |
| 二 | 上海行业新能源车配件价格分析 | | | 数据分析师 | |  |
|  |  | | |  | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 王琉立 | 工作年限 | 2 | | 专业 | 信息与计算科学 |
| 毕业学校 | 浙江财经大学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 北京行业数据清理 | | | 数据分析师 | |  |
| 二 | 上海行业新能源车配件价格分析 | | | 数据分析师 | |  |
|  |  | | |  | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 代和强 | 工作年限 | 8 | | 专业 | 计算机 |
| 毕业学校 | 杭州电子科技大学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 汽车定损系统 | | | 项目经理 | |  |
| 二 | 在线答题平台 | | | 项目经理 | |  |
| 三 | 城市管理系统 | | | 程序员 | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 徐君 | 工作年限 | 8 | | 专业 | 计算机 |
| 毕业学校 | 南京工程学院 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 保险经纪公司微商城 | | | 项目经理 | |  |
| 二 | 保险经纪公司电商平台 | | | 项目经理 | |  |
| 三 | 海峡定损APP | | | 程序员 | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 徐琮刚 | 工作年限 | 11 | | 专业 | 计算机科学与技术 |
| 毕业学校 | 西安电子科技大学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 浙江能源集团ERP实施项目 | | | 流程组长、技术顾问 | |  |
| 二 | 海康威视网上报销 | | | 开发顾问 | |  |
| 三 | 福建新龙马汽车股份有限公司ERP售前 | | | PI顾问 | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 贾健 | 工作年限 | 13 | | 专业 | 汽车维修 |
| 毕业学校 | 浙江省交通职业技术学院 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 中保协汽车保养负担指数采集与分析 | | | 项目负责人 | |  |
| 二 | 平安的EPC配件名称及关系的解析 | | | 项目负责人 | |  |
| 三 |  | | |  | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 莫立利 | 工作年限 | 10 | | 专业 | 电子商务 |
| 毕业学校 | 湖北大学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 中保协汽车保养负担指数采集与分析 | | | 业务数据专家 | |  |
| 二 | 平安的EPC配件名称及关系的解析 | | | 业务数据专家 | |  |
| 三 |  | | |  | |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 朱煌菲 | 工作年限 | 2 | | 专业 | 统计学 |
| 毕业学校 | 温州大学 | | | | | |
| 经 历 | | | | | | |
| 项目 | 参加过规划设计咨询项目的名称 | | | 担任何职 | | 备注 |
| 一 | 汽车配件二代系统 | | | 技术支持 | |  |
| 二 | 中保信上海定损系统 | | | 技术支持 | |  |
| 三 |  | | |  | |  |