

《数据分析及其在制造业应用》

——段方

教授 北京大学博士后

某世界 **500** 强企业大数据总设计师

1 概述

1.1 从泰勒经济学说起

1.1.1 科学管理之父

1.1.2 量化管理的发端

1.1.3 量化管理的价值和意义

1.2 数据分析基础——统计学

1.2.1 统计学的历史

1.2.2 统计学的价值

1.2.3 聚焦统计学

1.3 统计学的价值

1.3.1 收集数据

1.3.2 找出数理规律

1.3.3 从“拍脑袋”到“拍数据”

1.3.4 奠定数据分析的基础技术

1.4 制造业的统计学应用

1.4.1 数据渗透到企业的每个“毛孔”

1.4.2 大数据改变的思维模式

大数据的基本概念

用数据去“洞察”企业的每寸肌

肤 用“数理分析”构建企业

的“大脑 ”

1.4.3 如何发现“异常 ”？

降低次品率

降低成本

1.5 促进产业升级的价值和意义

1.5.1 从《美国工厂》说起

1.5.2 如何进行产业升级？

1.5.3 科学技术如何应用于企业内部管理

1.5.4 “一带一路”的战略价值——产业输出

1.6 统计学基础

1.6.1 从收集数据开始

1.6.2 数据分析方法的逐步演进

统计学

机器学习

深度学习

1.6.3 落地场景是关键

哪些环节收集数据

哪些环节使用数据分

析？如何尽快见到效

益？

1.7 【思考】如何从汽车制造业中利用数据分析进行产业升级？——丰田汽车的启示

2 统计学基础

2.1 统计学的产生和发展

2.1.1 统计学历史

古典统计学

近代统计学

现代统计学

2.1.2 历史上著名的统计学家

2.1.3 【思考】统计学与会计学间的联系

2.1.4 【例】算盘——最初的计算器

2.2 统计数据与统计学

2.2.1 数据研究过程

2.2.2 统计规律

2.2.3 【例】从“两圈半”说起

2.3 统计学的分科

2.3.1 统计分科

描述统计

推断统计

参数估计

假设检验

两者关系

2.3.2 应用领域

2.3.3 统计学与其它学科的关系

2.3.4 【例】网络上“被平均”的问题在哪里？

2.4 统计学的基本概念

2.4.1 总体和样本

2.4.2 标志和指标

2.5 【思考】统计学如何适配工业文明的发展

3 统计数据的描述

3.1 数据的收集

3.1.1 “巧妇难为无米之炊”

确立收集数据的意识

逐步建立数据思维

“没有人能够跟数据叫板 ”

3.1.2 直接渠道

抽样调查

统计报表

重点调查

典型调查

3.1.3 间接渠道

统计年鉴等

文献检索

3.1.4 数据的误差

抽样误差

非抽样误差

误差的控制

3.1.5 【例】附件-数据收集及质量控制

3.2 统计数据的整理

3.2.1 数据的预处理

3.2.2 统计分组

单变量值分组

组距分组

累积频数

3.2.3 数据分布的图示和类型

直方图

折线图

3.2.4 洛伦茨曲线

基尼系数

3.2.5 【例】数据质量的评估和控制

3.3 分布集中趋势的测度

3.3.1 众数

3.3.2 中位

数 3.3.3

均值

简单算术平均

加权算术平均

3.3.4 几何平均

3.3.5 众数、中位数和均值的比较

3.3.6 【例】从平均数到中位数的演进

3.4 分布离散程度的测度

3.4.1 极差

3.4.2 内距

3.4.3 方差和标准差

3.4.4 标准差的作用

3.4.5 离散系数

3.4.6 偏态系数和峰度系数

偏态系数

峰度系数

Subtop

ic

3.4.7 【例】中国和美国的“胖”的量化对比

3.5 统计图与统计表

3.5.1 统计表

3.5.2 统计图

茎叶图

箱线图

分布的形状与箱线图

3.6 【例】拿破仑真的是“矮子”吗？

4 数据分析在企业中应用

4.1 数据整理方法

4.1.1 从数据建模说起

4.1.2 二维报表的数据存储方法

4.1.3 数据库的存储方法

4.2 数据分析逻辑

4.2.1 问题的确定

4.2.2 围绕问题收集数据

4.2.3 分析方法的选择

4.2.4 分析方法具体运用

4.2.5 效果的后评估

4.3 企业的 KPI 举例

4.3.1 离职率

4.3.2 人工成本

4.3.3 销售数据

4.4 数据如何筛选

4.4.1 数据的口径一致问题

4.4.2 数据之间的血缘分析

4.4.3 关键变量选择

4.4.4 数据的计算过程

4.5 如何编写数据报告

4.5.1 如何学习写剧本

4.5.2 如何选定问题

4.5.3 如何设置叙事逻辑

4.5.4 组织相关数据

4.5.5 如何让别人懂得分析结果

4.6 条件概率和乘法公式的应用

4.6.1 条件概率

4.6.2 乘法定理

4.7 全概率公式和贝叶斯公式的应用

4.7.1 全概率公式与贝叶斯公式

4.7.2 事件的相互独立性

4.7.3 重要定理和结论

4.8 【例】 一些应用例题

5 制造业的数理统计应用案例

5.1 制造业统计学应用基础

5.1.1 “数据一思考，老板就发笑”

5.1.2 通用汽车的“4mm 工程”启示

5.1.3 制造业的管理水平评估角度——数据

5.2 制造业问题及对产销经营影响

5.2.1 制造业现场

5.2.2 现场改善的基础原则

5.2.3 依赖系统化科学方法

5.2.4 【例】丰田的现场管理经验

5.3 问题发掘与分析解决步骤/理念

5.3.1 现场问题的三个来源

5.3.2 解决现场问题的四个步骤

5.3.3 选择改善的主题

5.3.4 运用 I.E.手法

5.3.5 将改善落实于现场

5.3.6 运用稼动率分析技巧进行改善

5.3.7 工作抽样的进行程序

5.3.8 绩效分析法

5.3.9 现场问题分析技巧

MS

法

PS

法

5.3.10 分层图（Pareto）分析

5.4 现场品质控制问题

5.4.1 进料不良问题的分析与解决

5.4.2 进料检验方法的改善

外协厂商的管控

5.4.3 制程不良问题的分析与解决

5.4.4 Fool Proof 介绍

5.4.5 工序控制

工序能力

工序能力与公差

工序能力指数

5.4.6 设计工艺问题的分析与解决

5.5 进度交期问题的分析与解决

5.5.1 生产管理问题分析与解决

5.5.2 生产排程的两大重点

5.5.3 大日程计划量与生产批设定

5.5.4 粗略产能负荷管理基本格式

5.5.5 主排程的负荷展开主排程的负荷展开

5.5.6 细排程适用范围

5.5.7 MRP 模式的运用

5.5.8 物料可用量计算

5.6 现场成本问题

5.6.1 直接材料成本问题的分析与解决

5.6.2 价值工程和价值分析

5.6.3 价值的三要素

5.6.4 价值分析的 12 步工序

5.6.5 降低材料用量

5.6.6 直接人工成本问题分析与解决

5.6.7 流程线路图分析

5.7 现场效率问题

5.7.1 人均效率问题分析与解决

5.7.2 工作指导的六大步骤

5.7.3 工作教导的八个顺序

5.8 【思考】机器学习对制造业的影响

6 企业的大数据收集与整理

6.1 企业有哪些 IT 系统？

6.1.1 从 ERP 系统开始

6.1.2 OA 系统的数据价值

6.1.3 CRM 系统中的数据 “提纯”

6.1.4 财务数据的“洞察”

6.1.5 【调研】康明斯还有哪些数据？

6.2 内部能够产生哪些数据？

6.2.1 客户类的数据？

6.2.2 产品类的数据？

6.2.3 生产管理数据？

6.2.4 营销类数据？

6.2.5 人力资源数据？

6.2.6 财务数据？

6.2.7 非结构化数据

视频中的数据价值

地图数据的价值分析

6.2.8 【例】 GE 的发动机案例

6.3 如何从外部获取数据？

6.3.1 什么是数据的“维度”？

6.3.2 外部数据有哪些？

6.3.3 互联网数据的内容和特点

6.3.4 电信数据的内容和特点

6.3.5 车联网的数据呢？

6.3.6 其它的“数据金矿”

6.3.7 如何获取这些外部数据？

6.3.8 【例】 百度大数据的对外服务产品

6.4 如何将单点数据连接起来？

6.4.1 企业级数据模型

6.4.2 标准的数据字典

6.4.3 元数据的概念

6.4.4 数据治理的内容

6.5 如何连接客户的各种数据？

6.5.1 客户数据的内容和种类

6.5.2 车企能够掌握的数据

6.5.3 如何跨维度串联你客户的数据？

6.6 如何改进数据的质量？

6.6.1 什么是数据质量？

数据质量表现形式

是业务质量还是管理质量？

如何借助数据质量管理发动机产品质量？

6.6.2 数据质量的基本分析方法

血缘分析

影响分析

波动分析等方法

如何让人人参与数据质量工作中？

6.6.3 元数据的概念和基础

6.6.4 数据质量的关键点

为什么是 **60%**的工作量

不仅是技术问题

漫长而痛苦，却无法超越

6.7 【思考】汽车企业如何收集数据？

6.7.1 内部数据的梳理

6.7.2 外部数据的选择

6.7.3 如何产生更多的“金子”？

6.7.4 如何评估这些数据的质量？

6.7.5 如何打通这些数据的“联通”？

7 总结