

中级

单位、政府部门、IT 互联网企业、院校等的业务经营部门，从事数据分析与处理、经营行为分析、销售数据分析、数据分析报告撰写、高级运营、数据日常管理与维护、可视化应用等工作，根据业务发展需要，实现数据资源的整合和深度分析与应用。旅游大数据思维指导

旅游大数据业务指导 中级大数据工程师能够理清政府官方、主流 OTA、旅游协会等发布的旅游相关数据的来源、类型、采集方式等。中级大数据工程师能够遵守《旅游法》、《旅行社条例》处理旅游业务，并中级大数据工程师能结合大数据思维完成具体的业务工作。中级大数据工程师能够将旅游大数据灵活应用在两个或以上的具体旅游工作或业务。中级大数据工程师能够具体执行大数据与旅游相结合的落地路径。中级大数据工程师能够借用旅游大数据为上级领导的业务和管理决策提供参考素材。

旅游大数据咨询设计 中级大数据工程师能够独立对旅游大数据管理进行规划设计。中级大数据工程师能够依据已有的旅游业务场景和数据范围，构建数据思维导图、流程图。中级大数据工程师能够使用至少 2 种思维导图工具（如 Visio、ProcessOn 等），独立将具体业务进行分级细化，制定解决方案。中级大数据工程师能够独立分析旅游监管部门发布统计报告的内容，明确旅游统计的调查对象和调查范围。掌握旅游行业最新的发展变化，中级大数据工程师能够独立判断旅游行业政策、业务态势。

法规制度规范指导 熟悉并严格遵守旅游行业相关的重要法律法规。中级大数据工程师能够持续关注旅游行业具体业务最新的政策变化。熟悉并严格遵守大数据相关的重要法律法规。中级大数据工程师能够持续关注大数据方面最新的政策变化。

旅游数据收集与采集 **旅游数据采集规划** 中级大数据工程师能够独立按照正确的步骤和流程进行数据采集。不合格率不超过 20%。中级大数据工程师能够结合实际业务需求，独立制定旅游数据的采集计划，以指导具体数据采集工作。中级大数据工程师能够独立分析业务问题，明确所需要的数据以及数据来源、采集方法。中级大数据工程师能够正确设定和使用项目的数据采集标准，确保采集数据的标准化与合理性。中级大数据工程师能够独立根据项目需求，准确定位已有的内外部数据、旅游相关网站渠道等。

常见涉旅数据采集 中级大数据工程师能够借助常用的办公软件、在线工具独立进行快速高效的在线调研。中级大数据工程师能够借助常用的办公软件（如 Word、Excel 等）、数据相关软件，独立将线下纸质文件中的数据信息转化为电子数据。中级大数据工程师能够了解旅游企业内部业务软件相关的基本功中级大数据工程师能、核心业务逻辑和流程，筛选出旅游分析所需要的数据项。中级大数据工程师能够了解旅游监管软件相关的基本功中级大数据工程师能、核心管理逻辑和流程，筛选出旅游分析所需要的数据项。中级大数据工程师能够通过在线调查工具（如问卷星等）进行在线问卷设计；独立进行在线调研、调研统计。

互联网数据智中级大数据工程师能采集 了解互联网常见的数据分析网站，并中级大数据工程师能够根据业务需要，有效登录和正确查询这些网站的相关数据。了解涉旅数据常见的官方网站，并中级大数据工程师能够根据业务需要，有效且正确登录和查询这些网站的相关数据。中级大数据工程师能够借助至少 2 种网络爬虫工具（如 Excel 网页爬虫、后羿采集器、八爪鱼爬虫软件、爬山虎采集器等），独立且正确爬取与旅游业务相关的有效网络数据。了解大数据采集和传统数据采集的区别，中级大数据工程师能够掌握基本的采集方法。

旅游大数据管理与治理 **大数据编目管理** 中级大数据工程师能够对多渠道和手段采集的数据进行规范的数据归集与整理，归集成功率 90%。中级大数据工程师能够对不同类别的旅游数据资源进行简单的归类、编目，90%编目合理。中级大数据工程师能够对不同类别的旅游数据资源进行标签设计，标签创建成功率达到 90%。中级大数据工程师能够对已归集和编目的数据进行简单的数据集设置，以便标准化输出。中级大数据工程师能够在大数据管理平台，基于数据编目与管理进行基本的操作。

主数据管理 中级大数据工程师能够了解主数据与其他数据。中级大数据工程师能够评估数据源，对数据源进行标识、

比较和评估,了解主数据的唯一标识设定准则,数据源选择的标准,以及需求方的期望。中级大数据工程师能够合理地对归集的旅游数据进行的清洗。中级大数据工程师能够合理地对归集的旅游数据进行的重新编目与重构,形成数据工作台。

数据安全 中级大数据工程师能够合规合法地进行数据采集、分析与管理,符合数据安全、隐私的相关法律法规的要求。中级大数据工程师能够对不同的信息系统进行准确的信息安全等级保护定级。中级大数据工程师能够独立负责旅游信息系统二级等级保护所需的制度建设。中级大数据工程师能够自主参与配合信息系统二级等级保护制度以外其他建设和改造工作。中级大数据工程师能够参与旅游信息系统三级等级保护所需的制度建设。中级大数据工程师能够参与配合信息系统三级等级保护制度以外其他建设和改造工作。

数据质量管理 中级大数据工程师能够对旅游数据质量进行合理的评估。中级大数据工程师能够监控与评估数据质量改进的效果。

旅游大数据分析 中级大数据工程师能够掌握常见的数据统计指标,中级大数据工程师能够根据业务分析需求,选择正确的统计指标,并进行计算,计算正确率达 85%。中级大数据工程师能够独立从业务目标和实际旅游场景的角度,合理分析所获取与管理的大数据,分析结果合理性达到 80%。中级大数据工程师能够充分解读旅游企业经营、旅游政府监管等的数据分析报表,精准定位异常数据,分析出部分问题或结果,了解旅游相关业务中核心数据之间的关系。中级大数据工程师能够通过数据建模和设计,运用至少 2 种大数据分析工具(如 Excel、在线数据分析软件、SPSS 等)对旅游大数据资源进行统计、综合分析。

旅游大数据应用 中级大数据工程师能够基于旅游大数据业务场景,独立制作有针对性的数据分析报告。了解至少 5 个旅游大数据场景应用(如旅行服务(旅行社)、景区景点旅游大数据、酒店大数据、旅游假日大数据等)。中级大数据工程师能够独立通过对企业/政府自有数据以及外部数据的分析,深刻洞察旅游市场和旅游消费者,提供全方位的数据赋能,例如制定合理的市场营销策略、政府监管决策等,了解各个应用场景所需要的旅游大数据,以及数据要素之间的关联。中级大数据工程师能够将旅游大数据与其他产业、行业的大数据进行跨界融合,反哺赋能于旅游行业或其他行业。中级大数据工程师能够运用至少 1 种旅游大数据在线工具(如 Banber)制作数据大屏和数据报告。

大数据可视化 中级大数据工程师能够根据要展示的数据内容,独立选择合适、合规的可视化图表,合理性达到 85%。中级大数据工程师能够独立使用至少 2 种非程序式可视化工具(如 Excel、PPT、CBoard、百度 Sugar 等),对数据进行可视化解释。中级大数据工程师能够结合数据可视化,独立编写旅游大数据分析报告,报告数据的准确率达到 90%。中级大数据工程师能够完成数据统计图表化、数据结果展示化。主要面向企业、银行和非银行机构的投融资业务数据采集与分析、投融资分析和投融资风险管控等岗位,中级大数据工程师能利用大数据技术抓取结构化数据并进行数据处理、分析和数据可视化展示,业财税融合,从事投资分析、投资风险管理、融资分析和融资风险管理等工作。

数据采集与存储 中级大数据工程师能正确运用 Tushare 金融大数据中心采集债券市场利率、贷款利率和汇率等数据。中级大数据工程师能对支持 Python 数据接口的数据平台采集实时数据。中级大数据工程师能正确运用网络爬虫技术采集可投资领域的相关数据和信息。中级大数据工程师能正确将结构化数据存储到 MySQL 关系型数据库。中级大数据工程师能正确将非结构化数据存储到 MongoDB 非关系型数据库。

数据分析 中级大数据工程师能正确运用 Numpy 库进行数据计算。中级大数据工程师能正确运用 Lambda 库进行数据分析。中级大数据工程师能正确运用 Pandas 库进行数据分析。中级大数据工程师能正确运用正则表达式进行数据计算和分析。

数据可视化基础应用 中级大数据工程师能正确使用 Matplotlib 库生成折线图。中级大数据工程师能正确使用 Matplotlib 库生成柱状图。中级大数据工程师能正确使用 Matplotlib 库生成饼状图。中级大数据工程师能正确使用

Matplotlib 库生成雷达图。中级大数据工程师能正确使用 Matplotlib 库生成散点图。中级大数据工程师能正确使用 Matplotlib 库生成符合业务需求的图形投资分析 投资风险分析

中级大数据工程师能熟练利用大数据技术采集可投资领域的政策、税收等方面的信息数据。中级大数据工程师能根据采集的信息数据,结合企业内外部环境,科学分析政策风险、税务风险、行业风险、技术和环保风险等。中级大数据工程师能根据风险与报酬对等原则,合理评估企业或项目风险等级。

投资方案分析 中级大数据工程师能结合税收政策,运用收益法、市场法和成本法测算投资企业或项目的实体价值,科学分析企业或项目的投资方案。中级大数据工程师能运用净现值法、回收期法和内含报酬率法,科学分析独立投资方案。中级大数据工程师能运用共同年限法和等额年金法等判断标准,科学分析多个互斥投资方案。中级大数据工程师能结合税收政策,分别运用独立方案和互斥方案的判断标准,综合分析企业或项目投资。

投资可行性分析 中级大数据工程师能根据税收政策、现有的人力资源、财务资源和管理资源等企业内部信息,科学分析投资项目的可行性。中级大数据工程师能根据税收政策、企业或项目经营效益等数据和相关信息,定量分析企业或项目的投资收益。中级大数据工程师能根据国家法律法规规定和监管机构的要求,企业风险承受度和管控中级大数据工程师能力,科学分析企业或项目投资的可行性。中级大数据工程师能根据税收政策、财务指标和非财务指标,综合分析投资企业或项目,撰写企业或项目投资可行性分析报告。

投资风险管理 投资风险目标确定 中级大数据工程师能根据税收政策和企业或项目投资方案,科学确定投资成本目标和效益目标。中级大数据工程师能根据税收政策和企业或项目投资方案,合理确定投资效益目标。中级大数据工程师能根据企业或项目投资方案,科学确定投资安全警戒线目标。中级大数据工程师能根据企业或项目投资方案,科学确定投资工作进度目标。

投资风险管控制度制定 中级大数据工程师能根据企业战略目标,科学确定投资管理原则。中级大数据工程师能根据企业或项目投资方案,合理设置投资管理机构 and 风控机构。中级大数据工程师能根据企业或项目投资方案,确定各部门职责分工与工作流程。中级大数据工程师能根据企业或项目投资方案,制定投资风险管控制度。

投资风险防控 中级大数据工程师能科学诊断与评估企业风险现状,及时准确找出可中级大数据工程师能产生投资风险的关键环节。中级大数据工程师能根据投资风险的关键环节,科学制定投资风险管理目标。中级大数据工程师能根据投资风险管理目标,应用分散、备份等方法科学制定风险规避和风险转移等预案。中级大数据工程师能在风险发生后,及时制定风险应对优化方案。

融资分析 资本结构分析 中级大数据工程师能采用资产负债率 and 相关信息,准确分析企业或项目的资产负债水平。中级大数据工程师能采用产权比率、长期资本负债率 and 相关信息,准确分析企业或项目资金来源的渠道和水平。中级大数据工程师能采用股东权益比率 and 相关信息,准确分析企业或项目的自有资金和总资产的水平。中级大数据工程师能采用资本成本比较法和每股收 法,科学确定企业或项目最佳资本结构。

融资方案撰写 中级大数据工程师能从金融大数据中心采集和计算债券市场利率、股利增长率等数据,正确使用数据可视化工具生成融资需要的图形分析报表。中级大数据工程师能根据各债务融资贷款产品的优缺点、科学制定债务融资方案。中级大数据工程师能根据内部融资、优先股、普通股等股权融资方式的优缺点,科学制定股权融资方案。中级大数据工程师能根据行业 and 同类项目的市场利率 and 必要收益率等指标确定融资渠道,撰写规范的融资方案。中级大数据工程师能运用现金流量折现 and 股利折现等模型,准确分析比较不同融资方案。

还款计划制定 中级大数据工程师能根据信贷投放、政府补助、企业或项目内部现金流 and 经营等相关数据,科学分析企业或项目的还款中级大数据工程师能力。中级大数据工程师能根据企业经营预算、资金回笼变动 and 各类融资成本等数据,分析企业或项目的还款时间和金额,科学制定还款计划。中级大数据工程师能根据企业或项目的相关情况,科学选定还款方式。

融资风险管理 融资风险识别 中级大数据工程师

能根据企业或项目融资方案，正确识别融资方式风险。 中级大数据工程师能根据企业或项目融资方式，正确识别融资成本风险。 中级大数据工程师能根据企业或项目融资方案和税收政策，正确识别融资政策风险和融资税务风险。 融资风险管控制度制定 中级大数据工程师能根据企业战略目标，科学确定融资风险管理原则。 中级大数据工程师能根据企业或项目融资方案，合理设置融资管理机构和风控机构。 中级大数据工程师能根据企业或项目融资方案，确定各部门职责分工与工作流程。 中级大数据工程师能根据企业或项目融资方案，制定融资风险管控制度。 融资风险应对 中级大数据工程师能进行企业现状分析与评估，找出可中级大数据工程师能存在融资风险的关键环节。 中级大数据工程师能根据融资风险管控目标，划定风险控制线，建立预警机制。 中级大数据工程师能根据不同的风险类别，制定融资风险处理预案。 中级大数据工程师能在风险发生后，及时制定风险应对优化方案。益无差别点化分析报表。根据业务需求和 workflows，主要使用 Python、MySQL、Hadoop、Linux 和 Pandas 等工具，完成大数据平台搭建、大数据采集、大数据处理与存储、大数据分析挖掘、大数据可视化等工作。平台管理 软件安装 中级大数据工程师能够根据业务需求实现 Linux 系统的安装与基础配置中级大数据工程师能够根据业务需求实现 Python 集成开发环境的安装与基础配置中级大数据工程师能够根据业务需求实现 MySQL 的安装与环境配置中级大数据工程师能够根据业务需求实现 Linux 系统集群的搭建与基础配置中级大数据工程师能够根据业务需求实现 Hadoop 相关组件或框架的安装与基础配置 软件管理 中级大数据工程师能够根据业务需求实现 Linux 的状态监控与异常分析中级大数据工程师能够根据业务需求实现 Python 集成开发环境的模块安装与加载中级大数据工程师能够根据业务需求实现 PyCharm 的模块安装与加载中级大数据工程师能够根据业务需求实现 MySQL 的状态监控与权限管理 平台维护 中级大数据工程师能够根据业务需求定期监控各项运维数据中级大数据工程师能够根据业务需求进行平台用户权限管理中级大数据工程师能够根据业务需求在检测平台数据过程中通过数据分析发现业务问题并提出优化建议中级大数据工程师能够根据业务需求保障平台的安全、稳定运行大数据采集与存储 大数据采集 中级大数据工程师能够根据业务需求实现完整的外部数据采集方案设计中级大数据工程师能够根据业务需求基于 Request、lxml、re 等模块实现静态网页数据爬取中级大数据工程师能够根据业务需求基于 Selenium 等框架或逆向分析等相关技术实现动态网页数据爬取中级大数据工程师能够根据业务需求基于 Scrapy 等框架实现模块化的数据爬取大数据清洗中级大数据工程师能够根据业务需求基于 Python 实现数据清洗和初步加工中级大数据工程师能够根据业务需求基于 Python 的列表、元组、数据框等数据结构实现数据结构化处理中级大数据工程师能够根据业务需求基于 Python 实现数据类型转换中级大数据工程师能够根据业务需要优化现有的数据清洗流程，制定数据清洗的流程，识别和规避数据清洗过程中的风险 大数据存储与管理 中级大数据工程师能够根据业务需求实现将数据存储至 Excel、CSV、TXT 等文件中级大数据工程师能够根据业务需求实现将数据存储至 MySQL 数据库中级大数据工程师能够根据业务需求基于 Python 或 MySQL 实现数据关联、数据插入、数据修改等操作中级大数据工程师能够根据业务需求基于 Python 或 MySQL 等工具实现数据抽取大数据分析 中级大数据工程师能够根据业务需求基于 Python 实现数据异常值、缺失值和重复值的识别与处理中级大数据工程师能够根据业务需求基于 Python 实现数据变换、标准化、离散化等操作中级大数据工程师能够根据业务需求基于 Python 实现数据维规约、数值规约、特征构造等操作中级大数据工程师能够根据业务需求基于 Python 实现数据检索、数据排序等操作 大数据分析 中级大数据工程师能够根据业务需求基于 Python 实现数据质量分析、关联分析、特征分析等操作中级大数据工程师能够根据业务需求基于 Python 实现线性回归、岭回归等回归预测中级大数据工程师能够根据业务需求基于 Python 实现决策树、逻辑回归、支持向量机、贝叶斯等分类预测 中级大数据工程师能够根

据业务需求基于 Python 实现 K 均值聚类、密度聚类、期望最大化聚类等聚类分析中级大数据工程师能够根据业务需求基于 Python 实现回归模型、分类模型、聚类模型的模型评估及参数调优

大数据可视化 中级大数据工程师能够根据业务需求基于 Matplotlib、Seaborn、Pyecharts 等绘图模块或工具实现柱状图、散点图、饼图等图形绘制中级大数据工程师能够根据业务需求实现绘图参数的调整中级大数据工程师能够根据业务需求基于数据可视化结果总结出有效的分析结论,支撑市场决策中级大数据工程师能够根据业务需求基于数据可视化结果撰写相应的数据分析报告主要面向大数据行业、大数据处理与应用领域企事业单位的大数据爬虫工程师、大数据 ETL 工程师、大数据开发工程师、大数据应用开发工程师、报表开发工程师等岗位,中级大数据工程师能根据业务需求,利用各种数据系统工具,编写脚本程序完成工程化的离线或网络数据采集、数据处理、数据清洗等。

大数据工程化采集

网络数据采集 根据网络数据采集需求,中级大数据工程师能使用编程或爬虫框架,查询网络数据源信息,确定网络地址及网页格式,获得网页代码准确的标签信息。根据网页信息,中级大数据工程师能使用编程或爬虫框架,采集网络数据标签中相应数据并配置存储方式,并输出网络采集脚本。根据网络采集脚本及数据过滤需求,中级大数据工程师能使用编程或爬虫框架,编写网络采集过滤脚本,并输出网络采集数据。

离线数据采集 根据离线数据采集需求,中级大数据工程师能使用脚本方式,确定目标离线数据源格式,获取准确的格式信息。根据数据格式,中级大数据工程师能使用脚本方式,配置数据采集脚本进行结构化或半结构化离线数据采集,获得 离线采集数据并输出数据采集脚本。根据数据格式,中级大数据工程师能使用脚本方式,完成数据过滤脚本。根据编写的离线数据采集脚本,中级大数据工程师能使用脚本方式,测试离线数据采集脚本的采集、过滤、存储,获得目标数据并对脚本进行持久化操作。根据数据内容,中级大数据工程师能使用脚本创建存储采集数据的数据仓库结构并创建分区表及视图表。

作业调度 根据任务调度需求,中级大数据工程师能使用调度工具,配置数据采集模块与数据采集流程,进行采集作业调度并检查采集情况,并获得采集结果信息。根据联合调度需求,中级大数据工程师能使用调度工具,上传多个脚本文件,获得可调度执行的作业文件。根据作业文件类型,中级大数据工程师能使用调度工具,编写作业调度脚本,配置定时、触发信息,获得作业调度脚本。根据编写的作业调度脚本,中级大数据工程师能使用调度工具,调试作业调度效果。

大数据工程化处理

数据清洗 根据采集的原始数据集及完整性规则,中级大数据工程师能编写使用简单删除法或空值替代法处理离线数据集中缺省数据的脚本,获得完整的数据集。根据完整数据集及去重规则,中级大数据工程师能编写标记、删除离线数据集中重复字段的脚本,获得无干扰数据集。根据无干扰数据集及标准化规则,中级大数据工程师能编写归一处理离线数据集中不符合标准单位要求或给定结果集字段的脚本,获得标准化数据集。根据标准化数据集及可用性规则,中级大数据工程师能编写替换标记或删除离线数据集中不符合数据质量要求数据的脚本,获得可用数据集。根据可用数据集,中级大数据工程师能编写去除离线数据中无关字段的脚本,获得有效数据集。

数据计算 根据数据清洗后的数据集,中级大数据工程师能编写连接各数据表并根据处理需求进行数值计算、字符转换、时间计算等处理的脚本,获得正确处理后的多表数据集。根据多表数据集,中级大数据工程师能编写对多表数据进行字段合并、拆分等操作的脚本,获得字段对应的多表数据集。根据字段对应多表数据集,中级大数据工程师能编写连接、关联处理多表数据的脚本,获得关联整合数据集。根据关联整合数据集,中级大数据工程师能编写数据条件聚合、分组的脚本,获得关联计算数据集。

数据派生 根据关联计算的数据及数据特征,中级大数据工程师能编写脚本进行数据聚合结果分析并设计数据标签,创建各主题标签库。、 根据不同主题的标签,中级大数据工程师能编写脚本,对关联计算数据集中的数据,编写标签计算脚本并进行标签派生,获得含属性标签的数据集。根据已标签数据集,中级大数据工程师能

根据主题编写脚本划分数据表，创建业务主题数据表。根据主题数据表的业务需求，中级大数据工程师能使用脚本方式将同主题指标、维度、属性均关联的数据集进行数据组织，获得符合业务主题的宽表。

大数据工程化应用 OLAP 系统应用 根据数据表结构及特征，中级大数据工程师能编写脚本，基于现有的事实表，创建维度表，获得符合分析需求的基础数据表结构。根据事实表以及维度表，在指定项目中选择数据的加载方式，配置数据仓库中的事实表以及维度表的字段信息，获得符合业务分析需求的模型。根据配置完成的模型，以查看表格的形式或可视化的方式进行模型验证，对模型进行修改和调试获得修改后的模型。根据修改后的模型，创建并使用数据立方，编写条件查询语句进行数据查询，查询目标数据并存储。

搜索系统应用 根据数据结构及组织方式，中级大数据工程师能编写脚本、REST 请求命令，编写相应索引、类型及文档，获得匹配的索引结构。根据索引结构，中级大数据工程师能使用图形化工具或 REST 请求命令，将数据完整导入到分布式搜索组件，获得可全文检索的数据系统。根据搜索业务需求以及搜索组件中数据，中级大数据工程师能使用图形化工具或 REST 请求命令，进行简单数据查询，正确查询目标数据并存储。

报表系统应用 根据参考报表样式，中级大数据工程师能使用合适的可视化工具及脚本编程方法，配置所需展示离线数据源的连接信息，正确连通数据系统，并获得报表初始数据。根据报表数据及离线数据需求，中级大数据工程师能使用可视化工具及脚本编程方法，生成符合业务需求的可交互查询动态数据报表。根据数据报表，中级大数据工程师能使用可视化工具及脚本编程方法，中级大数据工程师能根据业务需求，完成可供决策支持的数据报表处理。中级大数据工程师能够根据业务需求，针对复杂数据资源，编写定制化的数据采集代码并应用数据采集工具，完成数据采集；中级大数据工程师能够根据大数据预处理方案，进行基本的描述性统计分析，针对数据存在的缺失值、重复值、异常值、一致性问题，采用相应的方法进行处理；中级大数据工程师能够对不同数据源的数据实体进行识别，并进行数据集成；中级大数据工程师能够根据大数据可视化方案，采用基本的大数据分析语言、工具和算法，对加工后的数据进行分析，并完成对离线和在线数据的可视化；中级大数据工程师能够根据大数据治理需求，在企业、社会、政府等的大数据治理过程中，参与数据质量管理体系的建设。

大数据采集 大数据采集知识与技中级大数据工程师能掌握 掌握大数据采集的概念，及其在大数据治理过程中的意义。掌握大数据采集过程中数据来源、数据规模、数据类型、数据存储方式等出现的难点。掌握三种以上大数据采集的方式（数据录入、文本文件导入、数据库文件导入、网页爬虫等）及其基本原理。掌握至少一种大数据采集通用工具（Python、Excel、Flume、Sqoop、八爪鱼等）的使用。

大数据 中级大数据工程师能够根据业务情况，进行大数据采集业务需求 采集方案制定 调研，明确数据采集需求。中级大数据工程师能够根据数据采集需求，进行数据资源的梳理。中级大数据工程师能够根据梳理好的数据资源，确定数据采集指标要求，并完成指标的说明及规范。中级大数据工程师能够根据梳理好的数据资源，确定数据采集的方式和工具。中级大数据工程师能够制定详细的大数据采集方案。

大数据采集实施 中级大数据工程师能够根据大数据采集方案，组织实施大数据采集任务。针对复杂数据资源，中级大数据工程师能够编写定制化的数据采集代码并应用数据采集工具，完成数据采集。中级大数据工程师能够指导初级人员完成基本的数据采集工作。中级大数据工程师能够根据数据质量及安全保障规范要求，保障数据质量及安全。

大数据预处理 大数据预处理知识与技中级大数据工程师能掌握 掌握大数据预处理的概念，及其在大数据治理过程中的意义。掌握大数据预处理的各个环节和内容。掌握大数据预处理各环节（数据清洗、数据集成、数据归约、数据变换等）的基本操作。掌握至少一种大数据开发的编程语言（Python、Java、存储过程等），解决基本的数据预处理问题。掌握至少一种大数据预处理通用工具（Excel、Python、Sql、存储过程、Kettle 等）的使用。

大数据预处理方案制定 中级大数据工程师能够根据业务需求，初步

明确大数据预处理需求。 中级大数据工程师能够进行基本的描述性统计分析，了解数据整体情况，明确需预处理的数据。 中级大数据工程师能够根据数据整体情况，初步选定合适的大数据预处理工具和方法。 中级大数据工程师能够初步制定大数据预处理方案。

大数据预处理实施 中级大数据工程师能够掌握不同类型数据文件（txt、csv、html、xml、xls 等）的加载。 中级大数据工程师能够针对数据存在的缺失值、重复值、异常值、一致性问题，采用相应的方法进行处理。 中级大数据工程师能够根据大数据预处理方案，对数据进行预处理。 中级大数据工程师能够对不同数据源的数据实体进行识别，并进行数据集成。

大数据可视化 中级大数据工程师能够掌握 掌握大数据可视化的概念 掌握可视化在大数据治理过程中的意义。 掌握大数据可视化的形式和应用。 掌握至少一种大数据可视化工具（Excel、Tableau、Python、R 等）的使用。 大数据可视化方案制定 中级大数据工程师能够根据业务需求，初步明确大数据可视化需求。 中级大数据工程师能够根据业务需求，进行描述性和探索性数据分析。 中级大数据工程师能够根据数据整体情况和大数据可视化需求，初步明确基本的数据可视化目标。 中级大数据工程师能够初步制定大数据可视化方案。

大数据可视化 中级大数据工程师能够运用合适的可视化工具，完成离线数据探索和分析结果的可视化。 中级大数据工程师能够运用合适的可视化工具，完成在线数据探索和分析结果的可视化。 中级大数据工程师能够基于数据可视化结果进行基本的分析和解释。 中级大数据工程师能够初步编写大数据可视化报告。

4 大数据治理 中级大数据工程师能够掌握 掌握大数据治理的概念和关键要素。 掌握大数据治理的对象、内容、框架和核心准则。 熟悉至少一种大数据治理的相关工具（ERWin、Excel、SVN、禅道等）、技术支撑和实施保障。 熟悉大数据治理的产业发展和应用场景。

数据质量管理 掌握大数据质量管理的概念和框架。 掌握影响数据质量的因素。 掌握数据质量评价模型和体系。 掌握数据质量过程控制。 掌握至少一种数据质量管理工具（Excel、存储过程、禅道等）的使用。 中级大数据工程师能够在企业、社会、政府等的大数据治理过程中，参与数据质量管理体系的建设。

大数据治理材料编制 中级大数据工程师能够根据业务需求，初步进行大数据治理的需求分析。 中级大数据工程师能够辅助制定数据标准管理的相关规范和制度。 中级大数据工程师能够辅助制定数据质量管理的相关规范和制度。 中级大数据工程师能够辅助进行大数据治理材料的编制。根据项目开发需求，完成大数据基础开发、大数据分析处理、大数据平台管理等工作。掌握大数据各个组件的开发与使用。

大数据平台组件维护管理 **大数据平台服务器管理** 按照大数据平台管理指导书，中级大数据工程师能够独立完成 Linux 文本处理的工作； 按照大数据平台管理指导书，中级大数据工程师能够独立完成 Linux 用户、权限管理工作； 按照大数据平台管理指导书，中级大数据工程师能够独立完成 Linux 硬件资源监控管理的工作； 按照大数据平台管理指导书，中级大数据工程师能够独立完成 Linux 软件安装的工作。

Hadoop 组件维护管理 按照大数据平台管理指导书，中级大数据工程师能够独立通过集群图形化管理应用程序启动 Hadoop 各组件； 按照大数据平台管理指导书，中级大数据工程师能够独立通过集群图形化管理应用程序停止 Hadoop 各组件； 按照大数据平台管理指导书，中级大数据工程师能够独立通过集群图形化管理应用程序查看 Hadoop 各组件的配置； 根据实际大数据平台管理要求，按照大数据平台管理指导书，中级大数据工程师能够独立通过集群图形化管理应用程序修改 Hadoop 各组件的配置。

Hive 组件维护管理 按照大数据平台管理指导书，中级大数据工程师能够独立通过集群图形化管理应用程序启动 Hive 组件； 按照大数据平台管理指导书，中级大数据工程师能够独立通过集群图形化管理应用程序停止 Hive 组件； 按照大数据平台管理指导书，中级大数据工程师能够独立通过集群图形化管理应用程序查看 Hive 组件的配置； 根据实际大数据平台管理要求，按照大数据平台管理指导书，中级大数据工程师能够独立通过集群图形化管理应用程序修改 Hive

组件的配置。按照大数据平台管理指导书,中级大数据工程师能够独立通过集群图形化管理应用程序启动 HBase 组件; HBase 组件维护管理 按照大数据平台管理指导书,中级大数据工程师能够独立通过集群图形化管理应用程序停止 HBase 组件; 按照大数据平台管理指导书,中级大数据工程师能够独立通过集群图形化管理应用程序查看 HBase 组件的配置; 根据实际大数据平台管理要求,按照大数据平台管理指导书,中级大数据工程师能够独立通过集群图形化管理应用程序修改 HBase 组件的配置。大数据组件开发 SQL 数据分析开发 按照大数据平台开发指导书,中级大数据工程师能够独立完成数据库、表的权限管理操作; 按照大数据平台开发指导书,中级大数据工程师能够独立完成表数据的 Join 连接查询操作; 按照大数据平台开发指导书,中级大数据工程师能够独立完成表的索引创建、使用等操作; 按照大数据平台开发指导书,中级大数据工程师能够独立完成视图的创建、更新、删除等操作; 按照大数据平台开发指导书,中级大数据工程师能够独立掌握函数的使用。 MapReduce 组件开发 理解 MapReduce 编程模型、运行机制; 按照大数据平台开发指导书,中级大数据工程师能够理解基础的 MapReduce 统计分析案例; 按照大数据平台开发指导书,中级大数据工程师能够独立提交 MapReduce 计算任务; 按照大数据平台开发指导书,中级大数据工程师能够独立查看 MapReduce 计算任务执行结果。 Hive 组件开发 理解 Hive 架构原理; 按照大数据平台开发指导书,中级大数据工程师能够掌握 Hive 复杂数据类型的开发; 按照大数据平台开发指导书,中级大数据工程师能够掌握 Hive 复杂查询的开发; 按照大数据平台开发指导书,中级大数据工程师能够掌握 Hive 函数的使用。 HBase 组件开发 理解 HBase 架构原理; 按照大数据平台开发指导书,中级大数据工程师能够以 SQL 方式完成表的创建、查看、禁用、删除等开发任务; 按照大数据平台开发指导书,中级大数据工程师能够以 SQL 方式完成表数据写入的开发任务; 按照大数据平台开发指导书,中级大数据工程师能够以 SQL 方式完成表数据查询的开发任务。 Spark 开发 理解 Spark 基本原理; 按照大数据平台开发指导书,中级大数据工程师能够掌握 Spark 开发规范; 按照大数据平台开发指导书,中级大数据工程师能够完成 RDD 的基础开发; 按照大数据平台开发指导书,中级大数据工程师能够独立提交 Spark 计算任务; 按照大数据平台开发指导书,中级大数据工程师能够独立提交 Spark 计算任务执行结果。大数据 ETL 工具开发 Sqoop 工具开发 按照大数据平台开发指导书,中级大数据工程师能够使用 Sqoop 完成关系型数据库数据的采集工作; 按照大数据平台开发指导书,中级大数据工程师能够使用 Sqoop 完成大数据平台数据的导出工作; 按照大数据平台开发指导书,中级大数据工程师能够掌握 Sqoop 全量数据处理、以及增量数据处理; 根据实际大数据数据采集的工作要求,按照大数据平台开发指导书,中级大数据工程师能够利用 Sqoop 完成关系型数据库数据采集工作。 Flume 工具开发 按照大数据平台开发指导书,中级大数据工程师能够熟练掌握 Flume 中 Spooling Directory Source、Kafka Source 等的开发; 按照大数据平台开发指导书,中级大数据工程师能够熟练掌握 Flume 中 Memory Channel 的高级配置用法; 按照大数据平台开发指导书,中级大数据工程师能够熟练掌握 Flume 中 HDFS Sink、Kafka Sink 等的开发; 根据实际大数据数据采集的工作要求,按照大数据平台开发指导书,中级大数据工程师能够利用 Flume 完成日志数据采集工作。 Kafka 工具开发 熟练掌握 Kafka 基础原理; 按照大数据平台开发指导书,中级大数据工程师能够熟练掌握 Kafka 的 Productor 、 Consumer 等操作命令; 根据实际大数据实时数据采集的工作要求,按照大数据平台开发指导书,中级大数据工程师能够利用 Kafka 完成实时数据采集工作。根据操作手册的要求,独立完成 spa 应用开发环境配置与部署,采用日志管理和配置管理等工具方法,完成复杂故障的识别、判断和处理,参与运维管理流程制定,掌握 SQL 语法和 python 编程语言,承担大数据应用部署和调试等工作任务。

1. 平台搭建 1. 1 云计算平台搭建 1 中级大数据工程师能根据操作文档搭

建 Openstack 私有云平台，安装 Nova、Swift、Glance、Dashboard 等组件。 2 中级大数据工程师能在专业人员指导下，协助完成私有云平台搭建过程中的排错工作。 3 中级大数据工程师能根据操作文档，正确使用公有云资源，如云服务器、对象存储服务、云网络服务、云安全服务、云数据库服务等。

1. 2 操作系统安装 2. 1.2. 2.2. 3.2. 4 中级大数据工程师能独立配置操作系统所需虚拟环境。中级大数据工程师能独立完成操作系统的安装与设置。中级大数据工程师能独立进行操作系统补丁安装。中级大数据工程师能同时使用图形化和命令行界面进行操作。

1. 3 Hadoop 系统部署 3. 1.3. 2.3. 3 3. 4 排查。 中级大数据工程师能根据操作手册进行 Hadoop 单节点部署。中级大数据工程师能根据操作手册进行 Hadoop 伪分布式部署。中级大数据工程师能根据操作手册进行 Hadoop 集群部署。中级大数据工程师能在搭建好的集群上运行自带测试用例，进行错误 2. 统调优 2. 1 运维管理 2. 1 中级大数据工程师能运用供应商提供的系统运维管理工具或者可视化界面，独立完成大数据架构下各个模块（如：HDFS、spark、配置信息）的日常运维管理操作，如：系统状态监测、系统日志收集、日常巡检等。2. 2 中级大数据工程师能运用 Shell 命令完成对各个节点的日常运维操作，如：运维脚本查询、存储使用情况、机器运行状态、网络防火墙配置。2. 3 中级大数据工程师能根据业务需求，运用数据库管理工具(Pl sql , navicat)，完成数据库日常监控和运维。2. 4 中级大数据工程师能协助高级技术支持人员梳理优化运维管理流程，补充系统运维操作手册等相关文档。

2. 2 故障管理 2. 2. 1 中级大数据工程师能运用日常运维实例和网络检索实例，定位日常简单故障，评估可中级大数据工程师能造成的影响范围，并及时上报。2. 2. 2 中级大数据工程师能运用系统故障处理的常用方法和工具，独立分析常见故障的原因，提出改进建议和方法措施。2. 2. 3 中级大数据工程师能根据软件提供的系统故障处理手册，运用故障诊断工具或系统自带的故障诊断命令，进行应急处理，保障系统的稳定运行。2. 2. 4 中级大数据工程师能在故障闭环后，根据故障现象、解决过程、核心原因、处理方案等多方面因素，编写相关故障解决文档。

2.3 日常巡检 2. 3. 1 中级大数据工程师能使用基础运维工具完成硬件资源信息的收集汇总并协助硬件管理人员完成周期性更新勘误。2. 3. 2 中级大数据工程师能使用基础运维工具完成软件部署信息的收集汇总并独立完成软件服务部署信息的周期性更新勘误。2. 3. 3 中级大数据工程师能使用基础运维工具完成网络资源分配信息的收集汇总，并协助网络资源管理人员完成对网络拓扑结构和资源 IP 的周期性更新勘误。2. 3. 4 中级大数据工程师能根据既有常用组织架构体系，并通过供应商所提供的配置管理工具，完成对具体岗位职务的日常运维工作。

3. 应用部署 3. 1 数据采集 3. 1 中级大数据工程师能运用大数据技术 (SparkStreaming、Kafka 等) 完成数据采集系统的搭建和基础配置。3. 2 中级大数据工程师能运用 Python 编写基础的爬虫脚本，完成对指定网页的数据爬取。3. 3 中级大数据工程师能运用 Flume 技术，完成大数据架构体系的系统日志收集。3. 4 中级大数据工程师能运用 ETL 工具(Kettle)，完成 excel/json/csv 等通用格式的数据采集任务。3. 5 中级大数据工程师能运用 Kettle 构建定时调度作业，完成对已有采集任务的自动优化。

3. 2 数据预处理 3. 2. 1 中级大数据工程师能运用 ETL 工具(Kettle)，完成数据转换，合并与拆分动作。3. 2. 2 中级大数据工程师能运用 ETL 工具(Kettle)，对关键字段完成数据脱敏。3. 2. 3 中级大数据工程师能运用 Python 语言，实现对已采集数据的清洗。3. 2. 4 中级大数据工程师能通过 ETL 工具(sqoop)，完成从关系数据库到 NOSQL 数据库的数据迁移。

3. 3 大数据存储 3. 3. 1 中级大数据工程师能运用大数据技术（如分布式文件系统 HDFS)维护、管理数据存储系统，优化存储资源利用率和计算效率。3. 3. 2 中级大数据工程师能独立完成 HBase 分布式集群环境部署。3. 3. 3 中级大数据工程师能运用 HBase Shell 完成日常节点维护和数据查询导出等维护动作。

3. 4 大数据处理 3. 4. 1 中级大数据工程师能运用 spark 技术，完成多中间件日志格式的解析和系统各种基础指标的的日常统计。3. 4. 2 中级大数据工程师

能运用 **Ne04j** 技术,完成对图形化数据的可视化输出。

3. 4. 3 中级大数据工程师能运用日常数据库技术,优化 ETL 工具流程,监控并维护例行数据 ETL 任务。

3. 4. 4 中级大数据工程师能运用标注工具 Labelme,完成对日常巡检图像数据的标注和预警。根据工业企业的业务需求,中级大数据工程师能够实现工业大数据应用编程,中级大数据工程师能够使用编程方式实现数据采集、分布式数据库 **Hbase** 应用、数据仓库 **hive** 应用、内存数据库 **redis** 应用,中级大数据工程师能够使用工具实现复杂功中级大数据工程师能的数据可视化。工业大数据应用编程

Hadoop 数据开发代码编写 中级大数据工程师能够根据 **Hadoop** 开发要求,正确安装配置 **Hadoop** 开发所需要的环境; 中级大数据工程师能够通过团队协作,根据工业数据开发项目的需求,使用面向数据处理的开发的技术实现简单代码编写; 中级大数据工程师能够初步完成具有海量数据处理中级大数据工程师能力的代码编写。代码调试 中级大数据工程师能够根据程序语法规则,初步完成错误语法的处理; 中级大数据工程师能够初步进行代码单步调试; 中级大数据工程师能够初步进行代码断点调试; 中级大数据工程师能够初步进行代码输入输出调试与异常处理调试。

分布式数据库 HBase 应用与开发 分布式数据库部署与设计 中级大数据工程师能够根据工业大数据业务场景的需求,规划设计 **hbase** 部署方案; 中级大数据工程师能够独立安装 **zookeeper** 软件; 中级大数据工程师能够独立安装分布式数据库系统; 中级大数据工程师能够根据业务需求,独立完成分布式数据库方案设计; 中级大数据工程师能够完成分布式数据库安装后的测试,验证软件安装部署的正确性。

分布式数据库的数据操作 中级大数据工程师能够根据业务需求,选择合适的过滤器进行数据的过滤操作; 中级大数据工程师能够根据业务需求,使用 **Java** 编程实现分布式数据库的数据操作; 中级大数据工程师能够根据业务需求,进行分布式数据库的批量处理。

分布式数据仓库 Hive 编程与优化 分布式数据仓库部署 中级大数据工程师能够根据工业大数据业务场景的需求,完成分布式数据仓库的部署规划设计; 中级大数据工程师能够按照部署规划设计,独立完成分布式数据仓库本地模式安装部署; 中级大数据工程师能够按照部署规划设计,独立完成分布式数据仓库远程模式安装部署; 中级大数据工程师能够完成分布式数据仓库安装后的测试,验证软件安装部署的正确性。

分布式数据仓库管理 中级大数据工程师能够根据业务需求,实现表的管理; 中级大数据工程师能够根据业务需求,选择合适的表类型存储数据; 中级大数据工程师能够通过查询语句向表中插入数据; 中级大数据工程师能够实现多表之间的连接查询。

分布式数据仓库性中级大数据工程师能优化 中级大数据工程师能够根据业务需求,选择合适的模式设计,实现数据的优化存储; 中级大数据工程师能够根据数据的特征,进行 **HQL** 语句的优化。

内存数据库 Redis 管理与应用 内存数据库部署与管理 中级大数据工程师能够根据工业大数据业务场景的需求,独立安装内存数据库; 中级大数据工程师能够根据业务的需求,实现 **Redis** 主从复制的配置; 中级大数据工程师能够根据业务的需求,实现 **Redis** 持久化的设置。

内存数据库的数据操作 中级大数据工程师能够实现 **Java** 到 **Redis** 服务器的连接; 中级大数据工程师能够通过连接池连接到 **Redis** 数据库; 中级大数据工程师能够使用 **java** 编程实现对多种数据类型的操作。

工业数据可视化 工业数据可视化设计 中级大数据工程师能够根据业务需求,完成工业数据可视化方案设计; 中级大数据工程师能够根据业务需求,对数据可视化设计方案进行优化。

工业数据可视化实现 中级大数据工程师能够使用数据可视化工具完成数据 **BI** 报表的制作; 中级大数据工程师能够基于数据可视化方案设计,使用数据可视化工具完成数据展示大屏的制作。

时序数据库存储与应用 时序数据库部署与管理 中级大数据工程师能够根据工业大数据业务场景的需求,独立完成时序数据库安装部署; 中级大数据工程师能够实现时序数据库的权限管理; 中级大数据工程师能够实现时序数据库的备份、还原、导出、导入的管理。

时序数据库的数据操作 中级大数据工程师能够根据业务需求,实现时

序数据的增加、查询、删除操作； 中级大数据工程师能够根据业务需求，实现时序数据的连续查询； 中级大数据工程师能够根据业务需求，实现保留策略的创建、查询、管理等操作。

时序数据库的应用实现 中级大数据工程师能够根据工业大数据业务场景完成 Telegraf 部署及设置； 中级大数据工程师能够根据工业大数据业务场景完成 Kapacitor 部署及设置； 中级大数据工程师能够根据工业大数据业务场景完成 Chronograf 部署及设置； 中级大数据工程师能够基于 TICK 技术栈搭建 DevOps 监控系统。根据业务需求，独立完成宏观、行业、上市企业以及金融产品相关接口数据采集、数据清洗、数据咨询服务、数据产品设计、轻量化的行业或企业微观数据分析等工作。

AI 自动化数据采集及清洗 数据规范 掌握数据源授权使用许可范围； 熟悉接口数据采集的技术标准和规范。 接口数据采集 掌握接口的请求方式（get, post, 请求头, 请求体等）； 掌握基本数据结构（json）； 熟悉基本网络传输协议（http、https 等）； 了解基本的 Python 语句。

文本数据处理 了解 OCR 识别技术的概念，NLP 自然语言处理技术的原理及使用场景； 中级大数据工程师能够使用文本数据采集工具对图片、 pdf 文件等数据进行处理。

数据清洗 掌握使用 python 语言进行数据去重的方法； 中级大数据工程师能够使用 python 语言的常见数据清洗手段对日期格式进行统一； 中级大数据工程师能够使用 python 语言对异常值及极端进行处理； 掌握使用 python 语言对空值及空格进行处理的方法。

数据库搭建及存储 数据库设计 中级大数据工程师能确定业务实施范围（业务用例）、用户行为（业务活动流程图）、输出 ER 图； 中级大数据工程师能够确保数据结构中级大数据工程师能够准确反映主题数据含义； 中级大数据工程师能够确保数据结构的完整性； 中级大数据工程师能够将 ER 模型转换为关系模型； 掌握根据关系模型确定数据库存储结构的方法。

数据库搭建 掌握 mysql 搭建数据库的操作及方法； 中级大数据工程师能够基于方案设计，进行数据库搭建； 掌握 mysql 数据库的操作，具备数据库的管理中级大数据工程师能力； 掌握 mysql 数据库的操作，具备数据库的运维中级大数据工程师能力。

接口数据入库存储 熟练掌握 sql 数据表的增、删、改、查操作； 中级大数据工程师能够使用 python 语言连接数据库； 中级大数据工程师能够使用 python 语言存储数据到数据库； 中级大数据工程师能够使用 python 语言在数据库中查找、提取及删除数据。

金融数据分析建模 银行业数据分析 了解银行业通用信贷风控业务场景； 掌握特征选择的方法，并筛选出对违约状态影响最显著的指标； 中级大数据工程师能够根据案例数据，建立模型； 掌握使用机器学习预测及监控风险的方法。

证券业数据应用 了解折线图、直方图、词云图等的使用场景； 掌握使用 python 语言根据市场情绪数据进行可视化绘图的方法； 中级大数据工程师能够根据可视化图表对市场情绪因子进行挖掘与分析； 中级大数据工程师能够根据证券市场微观结构，获取需要的因子； 根据选取因子，构建证券市场的分析模型； 中级大数据工程师能够根据所给案例数据，结合选取因子，对证券市场的情况进行预测。

保险业数据分析 根据保险产品信息，明确营销问题，理解获取数据的业务逻辑； 掌握 python 中数据挖掘相关算法，如：关联规则计算及聚类算法分析等，对营销数据进行分析； 中级大数据工程师能够使用 python 中函数绘制出变量相关性热力图及用户画像图表。

文本数据分析 掌握相关性分析的基本原理及方法； 中级大数据工程师能够使用 python 语言根据新闻词频数据及金融市场进行相关性分析。主要面向物流行业相关企业或各大中型企事业单位的物流部门、运营或服务部门，从事仓储数据分析指标、日常报表指标的制定，对仓储中入库、储存、拣货和分货等作业数据进行深入挖掘分析及监控，完成常规数据分析报告的撰写等工作，具有从数据角度及时发现物流仓储运营及设备运转过程中存在的问题，并提出仓储中关于作业流程、系统功中级大数据工程师能、设备配置等的中级大数据工程师能力。

仓储数 仓储数据 中级大数据工程师能够独立根据仓储数据化运营执行方案，在据分析指标与采集

方案制定 分析目标制定 熟悉仓储数据分析作用的基础上,明确仓储数据分析需求,高效开展相关工作。 中级大数据工程师能够独立与研发、规划、运营、售后等部门进行沟通,明确各部门对仓储数据分析的需求。 中级大数据工程师能够结合各部门对物流仓储数据分析的需求,运用物流仓储数据分析目标制定方法与技巧,制定出具有可行性的物流仓储数据分析目标。具有较强的逻辑分析中级大数据工程师能力以及跨部门沟通中级大数据工程师能力。

仓储数据分析指标制定 中级大数据工程师能够依据仓储数据分析目标,熟练使用仓储管理后台和数据分析辅助工具,提取现有数据指标。中级大数据工程师能够熟悉仓储常用数据指标及其内涵,使用 Excel 数据处理软件对提取到的现有数据指标进行科学、合理的分类整理。中级大数据工程师能够依据物流仓储数据分析目标和现有数据指标情况,独立运用数据分析指标制定方法与技巧,制定科学、合理的新数据指标,并扩展新指标维度。具备较强的数据指标分类、整理中级大数据工程师能力,以及创新数据指标的中级大数据工程师能力。

仓储数据采集与处理方案制定 中级大数据工程师能够在熟悉仓储数据采集工作流程的基础上,确定采集数据类型、采集渠道、采集时间区间、采集数据指标和采集口径。中级大数据工程师能够运用物流仓储数据采集方案制定方法与技巧,选择合适的采集工具、数据采集方法,制定物流仓储数据采集方案。中级大数据工程师能够运用仓储数据处理方案制定方法与技巧,制定仓储数据处理方案。掌握 Excel 数据处理工具、SQL 常用语句及可编程语言 Python、R,具备综合数据分析中级大数据工程师能力和较强的发现问题、解决问题的中级大数据工程师能力。

仓储数据挖掘分析 进货和入库作业分析 掌握进货和入库作业分析中各项指标的含义和作用,中级大数据工程师能够根据数据分析需求的目标,确定合适的分析指标、分析方法、数据展现形式及结果的呈现形式。 中级大数据工程师能根据进货和入库作业分析中特定分析需求,建立有效的数据分析指标,并中级大数据工程师能够确定新指标的准确计算方法,精确有效地反映问题的本质。 对数据具有高度敏感性,中级大数据工程师能够根据指标之间的关联性、指标的变化趋势推断分析指标变化的原因,分析并挖掘效率升高或降低的原因。 中级大数据工程师能够根据数据分析的目标,从进货和入库的运营数据挖掘发现业务中存在的问题,提出运营优化建议。 中级大数据工程师能够根据供应商供货及时性的度量指标,对供应商供货的及时性做出评价,准确识别出存在供货不及时的供应商。

储存作业分析 中级大数据工程师能够在理解储存作业设备及流程的基础上,设计并选择中级大数据工程师能够体现库存状态的指标。 中级大数据工程师能够运用描述性统计、趋势分析、同比环比分析以及频数频率分析等多种数据分析方法,对库存数据指标的原因及变化规律进行分析。 中级大数据工程师能够使用数据透视表和排序等数据处理方法,对库存进行 ABC 分析,利用折线图绘制 ABC 分类图,并根据周转速度和库存深度的差异,确定不同类别商品的出库和存储方式。 了解关联度计算方法和使用场景,中级大数据工程师能够熟练使用数据分析软件,根据历史订单计算商品关联度。

盘点作业分析 中级大数据工程师能够在熟悉盘点业务流程的基础上,设计并选择可以有效评估盘点作业的指标。 中级大数据工程师能够运用描述性统计、趋势分析、同比环比分析以及频数频率分析等多种数据分析方法,对盘点数据指标的变化规律及原因进行分析。 中级大数据工程师能够针对不同的盘点类型,抽取一定数量的商品实现盘点任务分配。 中级大数据工程师能够根据综合分析结果发现业务流程中的问题,并提出业务流程优化建议。

拣货作业分析 中级大数据工程师能够根据拣货作业设计并选择中级大数据工程师能够有效评估拣货作业总量和效率的指标。 中级大数据工程师能够运用多种数据分析方法,对拣货数据总量和效率指标的变化规律及原因进行分析。 理解订单分析的应用场景和范围,掌握拣货订单 EIQ 分析的概念和方法,中级大数据工程师能够使用 Excel 的数据透视表功中级大数据工程师能完成订单的 EIQ 分析。 中级大数据工程师能够根据综合分析结果发现业务流程中的问题,并提出业务流程优化建议。

分货作业分析 中级大数据工程师能够设计并选择

中级大数据工程师能够体现分货作业效率和效果的指标。中级大数据工程师能够结合实际分货作业场景，计算理论效率与实际效率的差距，并提出提升实际作业效率的措施。中级大数据工程师能够运用多种数据分析方法，对分货数据指标的变化规律及原因进行分析。中级大数据工程师能够通过分货数据分析，绘制仓库资源热力图，中级大数据工程师能够及时发现仓库运营问题，并提出相应的解决办法。

智中级大数据工程师能装备运行状态分析 中级大数据工程师能够独立完成设备运行状态数据的统计及整理，分析设备的各项运动机构的实际性中级大数据工程师能，洞察设备最终落地的实际运行效果。熟悉“总量、结构、速度”的异常分析方法，中级大数据工程师能够对一段时间内设备异常的发生情况绘制折线图，分析各种类、各等级异常数量占比，计算同期环比增长情况并进行趋势分析。中级大数据工程师能够独立通过数据变换计算并比较不同设备、不同仓库的设备单位工作量的异常发生情况。中级大数据工程师能够通过热力图分析仓库异常频发区域和不同种类异常分布情况，通过统计不同码点的异常数量判断码点运营状况。中级大数据工程师能够独立分析设备维度的异常数据，以设备为单位进行分析比较，并通过绘制箱线图，识别异常频发车辆。

智中级大数据工程师能装备可靠性分析 中级大数据工程师能够根据设备可靠性数据分析目标，针对智中级大数据工程师能设备在实际作业中产生的多种类型数据，设计适当的可靠性指标计算方法。中级大数据工程师能够通过 MTBF、MTTR、可靠度、失效率等指标评估物流设备的可靠性。中级大数据工程师能够依据智中级大数据工程师能设备作业数据，分析设备的可靠性水平，准确识别低可靠性的设备。中级大数据工程师能够通过分析仓库中智中级大数据工程师能设备的可靠性数据，识别高风险零部件，通过运用数据分析方法，配合相关部门进行异常定位。

仓储数据监控与分析报告撰写 仓储数据监控 熟练掌握仓储数据指标及其含义，依据仓储运营数据监控的要求，监控各类数据指标，发现变化趋势。中级大数据工程师能够运用仓储数据报表制作技巧，熟练使用表格处理软件的报表制作功中级大数据工程师能，制作数据报表。掌握异常数据识别方法，及时发现异常数据并进行分析和上报。具备良好的数据安全意识，认真细致的数据监控工作态度，以及较强的异常数据判断中级大数据工程师能力。

仓储数据分析报告撰写 中级大数据工程师能依据仓储阶段性数据分析目标，对仓储运营效果进行数据分析。中级大数据工程师能根据仓储大数据分析报告框架，运用数据分析报告撰写工具，高质高效完成仓储大数据分析报告的撰写。具备较强的文字表达中级大数据工程师能力、沟通中级大数据工程师能力以及系统化思维中级大数据工程师能力。

仓储运营优化建议提出 熟悉仓储运营的常见问题和核心关注点，从数据角度及时发现物流仓储项目在运营中存在的问题。掌握仓储运营常见问题解决方法，中级大数据工程师能够基于仓储运营优化经验和技巧，初步形成优化方案。中级大数据工程师能够依据数据分析结果，独立提出仓储运营中关于作业流程、系统功中级大数据工程师能等方面的建议。具有良好的异常数据洞察力和较强的发现、分析、解决问题的中级大数据工程师能力。

大数据咨询管理 大数据技术咨询处理 中级大数据工程师能够独立解答客户关于大数据基础理论相关的问题。中级大数据工程师能够独立解答客户关于大数据分析技术相关的问题。中级大数据工程师能够独立解答客户关于大数据可视化技术相关的问题。中级大数据工程师能够根据实际业务情况，输出大数据技术说明文档。

大数据平台咨询处理 中级大数据工程师能够根据业务需求，选择合适的大数据平台及方案。中级大数据工程师能够根据业务需求，选择合适的大数据存储计算产品。中级大数据工程师能够根据业务需求，选择合适的大数据可视化产品。中级大数据工程师能够解答关于大数据平台使用相关的问题。

大数据平台管理 中级大数据工程师能够根据大数据平台使用手册，开通和配置大数据计算平台。中级大数据工程师能够根据大数据平台使用手册，开通和配置大数据可视化服务。中级大数据工程师能够根据大数据平台使用手册，基于大数据计算平台进行基本的操作。中级大数据工程师能够根据大数

据平台使用手册，基于 大数据可视化服务进行基本的操作。 2、数据分析管理 数据质量管理 中级大数据工程师能够根据业务需求构建数据评估体系，独立对企业数据进行完整性评估。 中级大数据工程师能够根据业务需求构建数据评估体系，独立对企业数据进行规范性评估。 中级大数据工程师能够根据业务需求构建数据评估体系，独立对企业数据进行一致性评估。 中级大数据工程师能够根据业务需求构建数据评估体系，独立对企业数据进行准确性评估。13 分析报告撰写 中级大数据工程师能够根据数据分析可视化结果，独立 形成数据分析报告。 中级大数据工程师能够通过数据分析可视化结果，独立 得出分析结论，并体现在数据分析报告中。 中级大数据工程师能够根据数据分析可视化结果，独立 发现业务中存在的问题，并体现在数据分析 报告中。 中级大数据工程师能够根据数据分析可视化结果和发现 的问题，提出建议，并体现在数据分析报告 中。 数据挖掘分析 数据挖掘 中级大数据工程师能够独立编写较为复杂的数据提取语句，从业务线的底层数据表提取数据。 中级大数据工程师能够进行业务数据分析，通过监控及 分析，支持产品改进及新模式的探索。 中级大数据工程师能够构建基础业务数据分析体系，帮助确定各项业务数据指标。 中级大数据工程师能够基于数据挖掘分类算法，对业务 目标进行预测分析。 数据建模 中级大数据工程师能够构建用户数据模型，挖掘用户属性及用户喜好等需求，为个性化产品推荐提供支持。 中级大数据工程师能够构建产品、运营及活动用户行为 评估体系，通过数据分析对产品、运营、市 场提出建议并推动实施。 中级大数据工程师能够使用不同维度和方法，对模型效 果进行准确性评估。 中级大数据工程师能够通过调整参数，使模型效果不断 优化。 机器学习预测 中级大数据工程师能够基于机器学习平台，搭建基本的 机器学习流程。 中级大数据工程师能够基于机器学习平台，根据业务需 求进行分类、聚类分析。 中级大数据工程师能够基于机器学习平台，根据用户画像 分析，构建简单的推荐系统。 中级大数据工程师能够基于机器学习平台，根据用户行 为进行用户流失预测。掌握大数据思维，具备在工作中发现财务与 业务数据需求，并运用相关专业工具进行数据准备与清理、建模与分析、可视化 呈现的中级大数据工程师能力，具备运用数据库工具生成财务报表的中级大数据工程师能力。 财务与业务中的 数据应用 一需求 发掘 数据分析与业务和财 务的关系描述 了解数据分析的内涵 理解数据分析如何影响 业务 理解数据分析如何影响 财务 数据分析的应用框架 解析 掌握用 IMPACT 框架进 行数据分析的流程 具备财务人员必备的数据分析思维 掌握将通用业务与财务问题转换为数据分析问题 数据准备与清理 一 数据认知 数据存储技术运用 识别会计循环中的数据 使用与存储 了解关系型数据库的数据关系 了解数据字典的应用 数据提取、转换与加 载（ETL）运用 掌握数据提取方法 掌握数据转换方法 掌握数据加载方法 12 数据建模与评价 一 数据分析 数据分析预处理 实施测试计划，选取分 析方法 描述数据画像 缩减数据（规约） 实施数据分析 应用回归分析方法 应用分类分析方法 应用聚类分析方法 数据可视化呈现 一 结果分享 图表展示应用 定义分析结果的目标 受众 选取适合的展示图表 高阶展示应用 优化图表实现更佳呈现 效果 有效沟通及分享分析洞 见 财务人员的数据 库运用 一 实践案 例 从交易型数据中检索、汇总、生成财务报表 设计并创建关系型数据 库：交易型财务数据存储 从财务报表使用者的视 角设计并创建数据库：SQL 应 用 整合运用数据库技中级大数据工程师能与 初阶 Python 技中级大数据工程师能主要面向大数据平台高可用性部署实施、大数 据组件维护及使用、大数据平台维护及大数据平台优化等工作岗位。从事 Hadoop 高可用集群部署及配置、Hadoop 组件维护及使用、Hadoop 集群节点管理及维护、 大数据平台故障诊断及维护等工作，掌握高可用集群 Hadoop 配置方法，熟练使用 shell，掌握 Hadoop 集群节点及其服务的增删改等基本操作方法，熟悉常用系统性中级大数据工程师能诊断工具及集群监控管理工具，中级大数据工程师能独立排查和解决大数据平台常见问题， 优化集群性中级大数据工程师

能。大数据平台高可用部署 中级大数据工程师能配置集群主机之间时钟同步 中级大数据工程师能配置集群主机之间时钟同步 中级大数据工程师能配置高可用 Zookeeper 集群 中级大数据工程师能配置高可用 Zookeeper 集群 中级大数据工程师能配置高可用集群 Hadoop 文件参数 中级大数据工程师能正确配置高可用集群 Hadoop 文件参数 中级大数据工程师能启动初始化 JournalNode 服务 中级大数据工程师能启动 JournalNode 服务 中级大数据工程师能初始化 JournalNode 服务 中级大数据工程师能启动高可用集群 中级大数据工程师能启动高可用集群 中级大数据工程师能验证启动后的高可用集群 运行正常 中级大数据工程师能配置高可用集群 HDFS 自动切换模式 中级大数据工程师能正确配置高可用集群 HDFS 自动切换模式 大数据组件维护 中级大数据工程师能维护 HBase 组件 中级大数据工程师能删除 HBase 组件 中级大数据工程师能修改 HBase 组件 中级大数据工程师能维护 Hive 组件 中级大数据工程师能删除 Hive 组件 中级大数据工程师能修改 Hive 组件 中级大数据工程师能维护 Zookeeper 组件 中级大数据工程师能删除 Zookeeper 组件 中级大数据工程师能修改 Zookeeper 组件 中级大数据工程师能维护 Sqoop 组件 中级大数据工程师能删除 Sqoop 组件 中级大数据工程师能修改 Sqoop 组件 中级大数据工程师能维护 Flume 组件 中级大数据工程师能删除 Flume 组件 中级大数据工程师能修改 Flume 组件 3. 大数据平台维护 中级大数据工程师能增加和删除集群节点 中级大数据工程师能正确增加集群节点 中级大数据工程师能正确删除集群节点 中级大数据工程师能维护集群主机的内存和磁盘 中级大数据工程师能维护集群主机内存 中级大数据工程师能维护集群主机磁盘 中级大数据工程师能维护集群网络 中级大数据工程师能维护集群主机网络 中级大数据工程师能维护集群局域网络 中级大数据工程师能维护集群文件系统 中级大数据工程师能维护集群主机文件系统 中级大数据工程师能维护集群文件系统 中级大数据工程师能维护集群数据库 中级大数据工程师能维护集群数据库 4. 大数据平台优化 中级大数据工程师能优化 Linux 系统参数 中级大数据工程师能优化 Linux 系统的内存 中级大数据工程师能优化 Linux 系统网络 中级大数据工程师能优化 Linux 系统磁盘 中级大数据工程师能优化 Linux 文件系统 中级大数据工程师能优化 Linux 系统缓冲区 中级大数据工程师能优化 HDFS 配置参数 中级大数据工程师能调整优化 HDFS 配置文件中的参数 中级大数据工程师能优化 MapReduce 配置参数 中级大数据工程师能调整优化 MapReduce 配置文件中的参数 中级大数据工程师能优化 YARN 配置参数 中级大数据工程师能调整优化 YARN 配置文件中的参数 5. 大数据平台诊断 中级大数据工程师能诊断 HDFS 负载均衡问题 中级大数据工程师能对单节点 HDFS 负载均衡进行诊断 中级大数据工程师能对集群 HDFS 负载均衡进行诊断 中级大数据工程师能诊断 MapReduce 负载均衡问题 中级大数据工程师能对单节点 MapReduce 负载均衡进行诊断 中级大数据工程师能对集群 MapReduce 负载均衡进行诊断 中级大数据工程师能诊断集群节点故障问题 中级大数据工程师能使用集群日志对节点故障进行诊断 中级大数据工程师能使用集群告警信息诊断节点故障 中级大数据工程师能诊断集群组件服务故障问题 中级大数据工程师能使用集群日志诊断组件服务故障问题 中级大数据工程师能使用集群告警信息诊断组件服务故障问题"