**课程特色：**

本实时数据仓库项目与尚硅谷推出的离线数据仓库5.0项目相同，都依托于真实电商巨头的电商体系和实际业务场景，并参照离线数据仓库的建模理论全力打造。本课程在原有课程基础上有重大升级，在延续原有课程的实用与硬核的基础上，与离线数仓项目实现对接，并全面升级了实时数据仓库的建模理论，理论更经得起推敲，建模体系更完善，数据分析更高效！

本项目与离线数据仓库5.0项目可以进行无缝对接，采用相同的数据源以及同一套数据采集体系。使用Maxwell作为业务数据采集框架，实时监控业务数据库变动情况。

采用Flink作为核心的实时计算引擎，充分利用Flink的多种优秀特性，状态编程、自定义处理函数、CEP、CDC等，并对每一个指标的计算提供了基础API和Flink SQL的两种实现方式。在数仓的整体构建过程中，还充分研究探讨了其余多种大数据框架工具，例如Redis、HBase、ClickHouse等，利用每个框架的特点发挥最大作用。

通过本课程的学习，相信学员可以掌握企业实际开发过程中实时数据仓库的全搭建流程，深入理解Flink的高阶应用实例，掌握开发环节的多种框架技术。

**课程内容：**

搭建用户行为数据的实时数据采集框架Flume-Kafka-Flink；

采用Maxwell搭建业务数据采集体系；

使用Flink CDC结合配置表、HBase完成数仓DIM层的搭建；

参照数仓建模理论，分层搭建实时数据仓库，包括DWD层、DWS层、ADS层。

使用SpringBoot编写数据接口，结合Sugar完成实时指标结果数据的可视化展示；

**关键技术点：**

数仓架构深入讲解、离线架构与实时架构对比分析、SpringBoot项目搭建讲解、Flink CDC深入案例分析、使用侧输出流分流操作、Flink与HBase交互、Flink状态编程应用、Flink CEP循环模式匹配、双流join、旁路缓存、异步IO编码、ClickHouse多引擎讲解、Flink SQL、数据可视化接口编写、Flink实战优化方案。