目录20180810

[1 新版本 3](#_Toc521704192)

[1.1 246\_新特性介绍\_rec 3](#_Toc521704193)

[1.2 247\_新特性介绍-易用性：标准化SQL支持以及更合理的API\_rec 3](#_Toc521704194)

[1.3 248\_新特性介绍-高性能：让Spark作为编译器来运行\_rec 3](#_Toc521704195)

[1.4 249\_新特性介绍-智能化：Structured Streaming介绍\_rec 3](#_Toc521704196)

[1.5 250\_新特性介绍-Spark 1.x的Volcano Iterator Model技术缺陷分析 \_rec 3](#_Toc521704197)

[1.6 251\_新特性介绍-whole-stage code generation技术和vectorization技术\_rec 3](#_Toc521704198)

[1.7 252\_Spark 2.x与1.x对比以及分析、学习建议以及使用建议\_rec 3](#_Toc521704199)

[1.8 253\_课程环境搭建：虚拟机、CentOS、Hadoop、Spark等\_rec 3](#_Toc521704200)

[1.9 254\_开发环境搭建：Eclipse+Maven+Scala+Spark\_rec 3](#_Toc521704201)

[1.10 255\_SparkSession、Dataframe、Dataset开发入门(1)\_rec 3](#_Toc521704202)

[1.11 255\_SparkSession、Dataframe、Dataset开发入门(2)\_rec 3](#_Toc521704203)

[1.12 256\_Dataset开发详解-初步体验untypd操作案例：计算部门平均年龄与薪资 \_rec 3](#_Toc521704204)

[1.13 257\_Dataset开发详解-action操作：collect、count、foreach、reduce等\_rec 3](#_Toc521704205)

[1.14 258\_Dataset开发详解-基础操作：持久化、临时视图、ds与df互转换、写数据等\_rec 4](#_Toc521704206)

[1.15 259\_Dataset开发详解-typed操作：coalesce、repartition 不加密 4](#_Toc521704207)

[1.16 260\_Dataset开发详解-typed操作：distinct、dropDuplicates\_rec 4](#_Toc521704208)

[1.17 261\_Dataset开发详解-typed操作：except、filter、intersect\_rec 4](#_Toc521704209)

[1.18 262\_Dataset开发详解-typed操作：map、flatMap、mapPartitions\_rec 4](#_Toc521704210)

[1.19 263\_Dataset开发详解-typed操作：joinWith \_rec 4](#_Toc521704211)

[1.20 264\_Dataset开发详解-typed操作：sort\_rec 4](#_Toc521704212)

[1.21 265\_Dataset开发详解-typed操作：randomSplit、sample\_rec 4](#_Toc521704213)

[1.22 266\_Dataset开发详解-untyped操作：select、where、groupBy、agg、col、join \_rec 4](#_Toc521704214)

[1.23 267\_Dataset开发详解-聚合函数：avg、sum、max、min、count、countDistinct \_rec 4](#_Toc521704215)

[1.24 268\_Dataset开发详解-聚合函数：collect\_list、collect\_set\_rec 4](#_Toc521704216)

[1.25 269\_Dataset开发详解-其他常用函数\_rec 4](#_Toc521704217)

[1.26 270\_Structured Streaming：深入浅出的介绍\_rec 4](#_Toc521704218)

[1.27 271\_Structured Streaming：wordcount入门案例\_rec 5](#_Toc521704219)

[1.28 272\_Structured Streaming：编程模型\_rec 5](#_Toc521704220)

[1.29 273\_Structured Streaming：创建流式的dataset和dataframe\_rec 5](#_Toc521704221)

[1.30 274\_Structured Streaming：对流式的dataset和dataframe执行计算操作\_rec 5](#_Toc521704222)

[1.31 275\_Structured Streaming：output mode、sink以及foreach sink详解\_rec 5](#_Toc521704223)

[1.32 276\_Structured Streaming：管理streaming query\_rec 5](#_Toc521704224)

[1.33 277\_Structured Streaming：基于checkpoint的容错机制\_rec 5](#_Toc521704225)

[1.34 278\_Spark面试、简历中的项目编写以及实际生产环境的集群和资源配置等\_rec 5](#_Toc521704226)

# 新版本

## 246\_新特性介绍\_rec

## 247\_新特性介绍-易用性：标准化SQL支持以及更合理的API\_rec

## 248\_新特性介绍-高性能：让Spark作为编译器来运行\_rec

## 249\_新特性介绍-智能化：Structured Streaming介绍\_rec

## 250\_新特性介绍-Spark 1.x的Volcano Iterator Model技术缺陷分析 \_rec

## 251\_新特性介绍-whole-stage code generation技术和vectorization技术\_rec

## 252\_Spark 2.x与1.x对比以及分析、学习建议以及使用建议\_rec

## 253\_课程环境搭建：虚拟机、CentOS、Hadoop、Spark等\_rec

## 254\_开发环境搭建：Eclipse+Maven+Scala+Spark\_rec

## 255\_SparkSession、Dataframe、Dataset开发入门(1)\_rec

## 255\_SparkSession、Dataframe、Dataset开发入门(2)\_rec

## 256\_Dataset开发详解-初步体验untypd操作案例：计算部门平均年龄与薪资 \_rec

## 257\_Dataset开发详解-action操作：collect、count、foreach、reduce等\_rec

## 258\_Dataset开发详解-基础操作：持久化、临时视图、ds与df互转换、写数据等\_rec

## 259\_Dataset开发详解-typed操作：coalesce、repartition 不加密

## 260\_Dataset开发详解-typed操作：distinct、dropDuplicates\_rec

## 261\_Dataset开发详解-typed操作：except、filter、intersect\_rec

## 262\_Dataset开发详解-typed操作：map、flatMap、mapPartitions\_rec

## 263\_Dataset开发详解-typed操作：joinWith \_rec

## 264\_Dataset开发详解-typed操作：sort\_rec

## 265\_Dataset开发详解-typed操作：randomSplit、sample\_rec

## 266\_Dataset开发详解-untyped操作：select、where、groupBy、agg、col、join \_rec

## 267\_Dataset开发详解-聚合函数：avg、sum、max、min、count、countDistinct \_rec

## 268\_Dataset开发详解-聚合函数：collect\_list、collect\_set\_rec

## 269\_Dataset开发详解-其他常用函数\_rec

## 270\_Structured Streaming：深入浅出的介绍\_rec

## 271\_Structured Streaming：wordcount入门案例\_rec

## 272\_Structured Streaming：编程模型\_rec

## 273\_Structured Streaming：创建流式的dataset和dataframe\_rec

## 274\_Structured Streaming：对流式的dataset和dataframe执行计算操作\_rec

## 275\_Structured Streaming：output mode、sink以及foreach sink详解\_rec

## 276\_Structured Streaming：管理streaming query\_rec

## 277\_Structured Streaming：基于checkpoint的容错机制\_rec

## 278\_Spark面试、简历中的项目编写以及实际生产环境的集群和资源配置等\_rec