



Enterprise Computing Solutions - Education Services

## TRAINING OFFERING

---

**You can reach us at:**

Arrow ECS, Woluwedal 30, 1932 Sint-Stevens-Woluwe

Email: [education.ecs.benelux@arrow.com](mailto:education.ecs.benelux@arrow.com)

Phone: +32 2 332 19 57



# IBM InfoSphere QualityStage Essentials v11.5

CODE:	LENGTH:	PRICE:
KM213G	32 Hours (4 days)	€3,120.00

## Description

This course teaches how to build QualityStage parallel jobs that investigate, standardize, match, and consolidate data records. Students will gain experience by building an application that combines customer data from three source systems into a single master customer record.

## Objectives

- List the common data quality contaminants
- Describe each of the following processes: §Investigation §Standardization §Match §Survivorship
- Describe QualityStage architecture
- Describe QualityStage clients and their functions
- Import metadata
- Build and run DataStage/QualityStage jobs, review results
- Build Investigate jobs
- Use Character Discrete, Concatenate, and Word Investigations to analyze data fields
- Describe the Standardize stage
- Identify Rule Sets
- Build jobs using the Standardize stage
- Interpret standardization results
- Investigate unhandled data and patterns
- Build a QualityStage job to identify matching records
- Apply multiple Match passes to increase efficiency
- Interpret and improve match results
- Build a QualityStage Survive job that will consolidate matched records into a single master record
- Build a single job to match data using a Two-Source match

## Audience

- Data Analysts responsible for data quality using QualityStage
- Data Quality Architects
- Data Cleansing Developers

## Prerequisites

Participants should have:

- Familiarity with the Windows operating system
- Familiarity with a text editor Helpful, but not required, would be some understanding of elementary statistics principles such as weighted averages and probability.

## Programme

1. Data Quality Issues• Listing the common data quality contaminants• Describing data quality processes2. QualityStage Overview• Describing QualityStage architecture• Describing QualityStage clients and their functions3. Developing with QualityStage• Importing metadata• Building DataStage/QualityStage Jobs• Running jobs• Reviewing results4. Investigate• Building Investigate jobs• Using Character Discrete, Concatenate, and Word Investigations to analyze data fields• Reviewing results5. Standardize• Describing the Standardize stage• Identifying Rule Sets• Building jobs using the Standardize stage• Interpreting standardize results• Investigating unhandled data and patterns6. Match• Building a QualityStage job to identify matching records• Applying multiple Match passes to increase efficiency• Interpreting and improving Match results7. Survive• Building a QualityStage survive job that will consolidate matched records into a single master record8. Two-Source Match• Building a QualityStage job to match data using a reference match

## Further Information

Prior to enrolling, IBM Employees must follow their Division/Department processes to obtain approval to attend this public training class. Failure to follow Division/Department approval processes may result in the IBM Employee being personally responsible for the class charges.

GBS practitioners that use the EViTA system for requesting external training should use that same process for this course. Go to the EViTA site to start this process: <http://w3.ibm.com/services/gbs/evita/BCSVTEnr1.nsf>

Once you enroll in a GTP class, you will receive a confirmation letter that should show:

The current GTP list price

The 20% discounted price available to IBMers. This is the price you will be invoiced for the class.

## **Session Dates**

On request. Please [Contact Us](#)

## **Additional Information**

This training is also available as onsite training. Please contact us to find out more.