



DPC: 25/30476603 DC

BSI Group Headquarters

389 Chiswick High Road London W4 4AL

Tel: +44 (0)20 8996 9000

Fax: +44 (0)20 8996 7400

www.bsigroup.com

Date: 09 September 2025

Origin: International

Latest date for receipt of comments: 02 November 2025

Project No. 2023/01500

Responsible committee: MS/6 Methodologies for business process improvement using statistical methods

Interested committees:

Title: Draft BS ISO 24481 Statistical methods for implementation of Six Sigma- Exploratory data analysis. Part 1: General methodology.

Please notify admin.start@bsigroup.com if you are aware of any keywords that might assist in clarifying or identifying the standard or of the content of this standard:

- i) has any issues related to 3rd party IPR, patent or copyright
- ii) affects other national standard(s)
- iii) requires additional national guidance or information

WARNING: THIS IS A DRAFT AND MUST NOT BE REGARDED OR USED AS A BRITISH STANDARD. THIS DRAFT IS NOT CURRENT BEYOND 02 November 2025

This draft is issued to allow comments from interested parties; all comments will be given consideration prior to publication. No acknowledgement will normally be sent. **See overleaf for information on the submission of comments.**

No copying is allowed, in any form, without prior written permission from BSI except as permitted under the Copyright, Designs and Patent Act 1988 or for circulation within a nominating organization for briefing purposes. Electronic circulation is limited to dissemination by e-mail within such an organization by committee members.

Further copies of this draft may be purchased from BSI Shop <http://shop.bsigroup.com> or from BSI Customer Services, Tel: +44(0) 20 8996 9001 or email cservices@bsigroup.com. British, International and foreign standards are also available from BSI Customer Services.

Information on the co-operating organizations represented on the committees referenced above may be obtained from <https://standardsdevelopment.bsigroup.com/>.

E-mail: admin.start@bsigroup.com

Introduction

This draft standard is based on national and international discussions. Your comments on this draft are invited and will assist in the preparation of the consequent standard.

For international standards, comments will be reviewed by the relevant UK national committee before sending the consensus UK vote and comments to the international committee, which will then decide appropriate action. If the international standard is approved, it is usual for the text to be published as a British Standard.

For national standards, comments will be reviewed by the relevant UK national committee and the resulting standards published as a British Standard.

UK Vote

Please indicate whether you consider the UK should submit a negative (with supporting technical reasons) or positive vote on this draft. Please indicate if you are aware of any reason why this draft standard should not be published as a British Standard.

Submission of Comments

- **Annotated drafts are not acceptable and will be rejected.**
- All comments should be submitted online at <https://standardsdevelopment.bsigroup.com/>. You will need to register in order to comment.

Template for comments and secretariat observations

Date: xx/xx/20xx	Document: ISO/DIS xxxx
------------------	------------------------

1	2	(3)	4	5	(6)	7
MB	Clause No./ Subclause No./Annex (e.g. 3.1)	Paragraph/ Figure/Table/Note	Type of comment	Comment (justification for change) by the MB	Proposed change by the MB	Secretariat observations on each comment submitted
	3.1	Definition 1	ed	Definition is ambiguous and needs clarifying.	Amend to read '...so that the mains connector to which no connection...'	
	6.4	Paragraph 2	te	The use of the UV photometer as an alternative cannot be supported as serious problems have been encountered in its use in the UK.	Delete reference to UV photometer.	



DRAFT International Standard

ISO/DIS 24481

Statistical methods for implementation of Six Sigma- Exploratory data analysis — Part 1: General methodology

ICS: 03.120.30

ISO/TC 69/SC 7

Secretariat: **SAC**

Voting begins on:
2025-09-09

Voting terminates on:
2025-12-02

This document has not been edited by the ISO Central Secretariat.

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENTS AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

Reference number
ISO/DIS 24481:2025(en)

Full document details: 25/30476603 DC

© ISO 2025  **Documents**
Document Sharing Hub



COPYRIGHT PROTECTED DOCUMENT

© ISO 2025

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Symbols and abbreviated terms	3
5 Basic concepts	3
5.1 General	3
5.2 Relation and difference between data visualization and EDA	4
5.3 Principles and related steps for conducting EDA	4
6 Tools of EDA from different perspectives	6
6.1 General	6
6.2 Certain tools for the different types of data	6
6.2.1 Discrete data	6
6.2.2 Continuous data	6
6.3 Sorted according to their objectives	7
6.4 EDA tools used in each phase of DMAIC	7
7 Emphases and cautions for certain tools and other aspects of EDA	8
7.1 General	8
7.2 Importance of checking the metadata	8
7.3 Limitation of descriptive statistics	9
7.4 Cautions about some graphical methods	10
7.4.1 Use graphical methods efficiently	10
7.4.2 Bar charts or dot charts instead of Pie charts	11
7.4.3 Avoid misleading graphs and note the data Inking issue	11
7.4.4 Causation is more important than correlation	11
7.4.5 The relationship between EDA and CDA or other tools	11
7.4.6 Suitable data analysis software especially for EDA	12
Annex A (informative) Comparison of data visualization and EDA from different aspects	13
Annex B (informative) Some EDA principles in Six Sigma and quality improvement projects	14
Bibliography	16

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

ISO XXXX-1 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 7, *Applications of statistical and related techniques for the implementation of Six Sigma*.

ISO XXXX consists of the following parts, under the general title Statistical methods for implementation of Six Sigma — Exploratory data analysis—:

- *Part 1: General methodology*
- *Part 2: Tools and techniques for one or two-dimensional data*
- *Part 3: Tools and techniques for large datasets and high-dimensional data*

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Like many other sciences and technologies, a core idea of Six Sigma is data-driven decision making, that is, exploiting the data from measurements or simulations at various points in the life cycle of a product or certain service. It is widely accepted that statistics plays a very important role in such fields. In any situation where statistics is applied, the analyst will follow a process, more or less formal, to reach findings, recommendations, and actions based on the data. There are two phases in this process:

1. Exploratory Data Analysis (EDA)
2. Confirmatory Data Analysis (CDA)

In 1970s, EDA was promoted by John Tukey to encourage statisticians to explore the data, and possibly formulate hypotheses. Generally speaking, the emphasis in EDA is on letting data talk by themselves or hypothesis generation. In EDA efforts, the analyst searches for clues in the data that help identify theories about underlying behaviours. In contrast, the focus of CDA is hypothesis testing and inference. CDA consists of confirming these theories and behaviours. CDA follows EDA, and together they make up statistical modelling during data analysis.

Since the rapid development of big data and data science, there is a clear need to evolve the mechanics of Six Sigma both to accommodate the greater availability of data and to address the fact that, historically, approaches to analysing data were overly concerned with hypothesis testing, to the detriment of the hypothesis generation and discovery needed for improvement. Six Sigma teams often find warehouses of data that are relevant to their efforts. Discovery is largely supported by the generation of hypotheses—conjectures about relationships and causality. Today's Six Sigma teams, and data analysts in the business world in general, are often trained with a heavy emphasis on hypothesis testing, with comparatively little emphasis given to hypothesis generation and discovery. They are often hampered in their problem-solving and improvement efforts by the inability to exploit exploratory methods, which could enable them to make more rapid progress, often with less effort.

Though the concept of EDA has evolved and is still evolving, the main idea of it is unchanged. In most data analysis process, EDA is an approach for data analysis that employs a variety of techniques (mostly graphical) to

1. maximize insight into a data set;
2. identify data issues (e.g., outliers, missing values, anomalies);
3. uncover underlying structure, patterns and relationship;
4. extract important variables;
5. determine optimal factor settings; and
6. enhance problem-solving, root cause analysis, and continuous improvement efforts in process optimization.

In recent times, we have seen incredible advances in visualization methods which is very important and useful to EDA, supported by phenomenal increases in computing power. Some software packages including MATLAB¹⁾, SAS-JMP¹⁾, Minitab¹⁾, Tableau, Python and R are well developed to incorporate such methods. With these preparations, it is believed that the approaches now allowed by these methods as well as EDA are underutilized in current Six Sigma practice.

1) MATLAB is the trade name of a product supplied by Matlab Inc. JMP is the trade name of a product supplied by SAS Institute Inc. Minitab is the trade name of a product supplied by Minitab LLC. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these products.

