Eastern Economy Edition

THIRD EDITION

DATA WAREHOUSING

Concepts, Techniques, Products and Applications

C.S.R. Prabhu

PHI

DATA WAREHOUSING

Concepts, Techniques, Products and Applications

DATA WAREHOUSING

Concepts, Techniques, Products and Applications

THIRD EDITION

C.S.R. PRABHU
Deputy Director General
National Informatics Centre
Hyderabad

PHI Learning Private Limited
New Delhi - 110001
2008

DATA WAREHOUSING—Concepts, Techniques, Products and Applications, 3rd ed.
C.S.R. Prabhu

© 2008 by PHI Learning Private Limited, Delhi. All rights reserved. No part of this book may be reproduced in any form, by mimeograph or any other means, without permission in writing from the publisher.

ISBN-978-81-203-3627-8

The export rights of this book are vested solely with the publisher.

Fifteenth Printing (Third Edition) ... ... January, 2015

Published by Asoke K. Ghosh, PHI Learning Private Limited, Rimjhim House, 111, Patparganj Industrial Estate, Delhi-110092 and Printed by Rajkamal Electric Press, Plot No. 2, Phase IV, HSIDC, Kundli-131028, Sonepat, Haryana.

Contents

Preface

Data Warehousing, Online Analytical Processing (OLAP) and Data Mining represent some of the latest trends in computing environment and information technology applications to large-scale processing and analysis of data.

Data warehouses along with OLAP tools are being increasingly developed to analyse historical or time-series data to identify past patterns or trends which may be useful in forecasting the future. In addition, the unknown associations and relationships of data items in data-intensive areas can be found out by using various data mining techniques, resulting in knowledge discovery in the data.

This book clearly describes data warehousing concepts at an introductory level. The currently available software tools used for design and development of data warehouses are compared and contrasted. For readers interested in developing data warehouses for their organizations, the text offers practical guidelines that can help architect data warehouses. For applications in government sectors such as agriculture, rural development, finance and commerce, the book identifies the broad directions and enumerates the steps involved in developing data warehouses for decision support purposes. The real data warehouse implementations already performed successfully in government sectors are presented as Case Studies.

The book in the present third edition (after a successful second edition) is organized into five chapters with some of the chapters enhanced and enriched. Chapter 1 highlights the fundamental concepts of data warehousing and data marts, besides covering issues related to creation, maintenance and refreshing of data marts. Chapter 2 analyses the design methodology of OLAP tools and how they differ from online transaction processing (OLTP), star schema, snow flake schema and also gives an overview of various currently available OLAP tools. With open source platforms becoming popular, open source data warehousing and OLAP tools (such as Pentaho) have been added in this chapter. Chapter 3 elucidates with suitable examples the concepts and techniques used in data mining. In addition, in Chapters 2 and 3 some highlights of open source OLAP and data mining tools are also added. In Chapter 4 the practical approaches to selection of hardware, software, communication and data management systems

for building real data warehouses are discussed. Chapter 5 enumerates the potential applications of data warehousing and data mining in various government sectors.

The case studies of data warehouse implementations for the Tamil Nadu and Andhra Pradesh governments and Export Processing Zones (EPZs) presented in this book offer detailed insights into practical knowledge required for such projects. The case studies of data warehousing in World Bank, Hewlett-Packard and Levi Strauss indicate the efforts made in banking and business sectors in data warehouse implementation. In this edition four new case studies have been added of which three relate to banking and retail sectors. The fourth one is on a highly available data warehouse.

The book aims at software professionals and database practitioners who are interested in developing an understanding of concepts and techniques of data warehousing. Engineering students pursuing M.Tech courses in computer science and those studying computer applications (MCA) will also immensely benefit from this single point coverage of the basics and practical guidelines to building data warehouses.

The book endeavours to provide sound and practical inputs to understand, plan and build data warehouse applications. It is hoped that this text will serve the needs of students and teachers for introductory courses in data warehousing. Any suggestions for improvements to the text will be gratefully acknowledged.

C.S.R. Prabhu

Acknowledgements

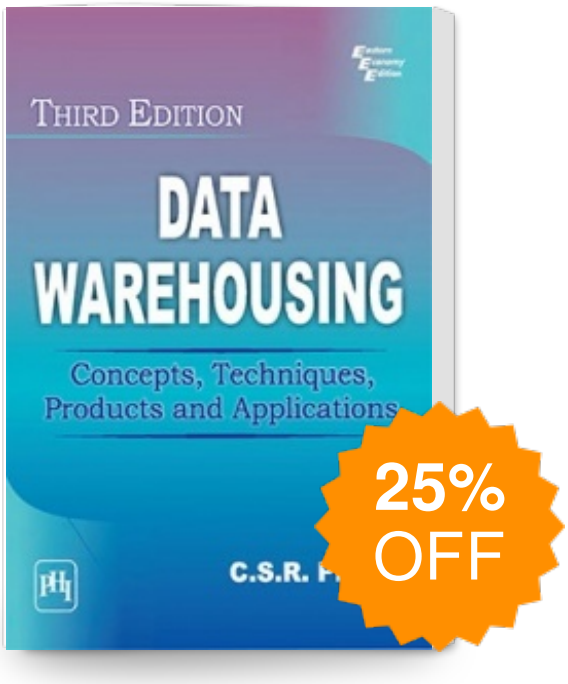
I am deeply indebted to the Director General, National Informatics Centre (NIC), New Delhi, for giving me the permission to write this book. I thank Microsoft (India) for permitting me to cover case studies on Hewlett-Packard data warehouse and World Bank data warehouse. I wish to thank and acknowledge McGraw-Hill for permission to use a few diagrams and some product descriptions from their book *Data Warehousing, Data Mining and OLAP*. I also sincerely thank and acknowledge the authors of the Web site in public domain on the data warehousing in Levi Strauss built by ArsDigita Corporation.

Numerous companies have also permitted me to cover their product descriptions. I am grateful to all of them. I thank Mr. V.S. Raghunathan of NIC, Chennai, for the case study report on Export Processing Zones (EPZs). I thank Mr. B.M. Rao, Mr. Guruprasad and Mr. D. Laxminarayana of NIC, Hyderabad for assistance in preparing case studies on A.P. Government Data Warehouses. I also thank Mrs. Annapurna, Mr. Shravan Veer and Mrs. Ratna Sudhakar of NIC, Hyderabad in preparing the manuscript.

Finally, I thank my Publishers, Prentice-Hall of India, for the careful processing of the manuscript both at the editorial and production stages.

C.S.R. Prabhu

Data Warehousing : Concepts, Techniques, Products And Applications



Publisher : PHI Learning ISBN : 9788120336278 Author : C S R Prabhu

Type the URL : http://www.kopykitab.com/product/7772

Get this eBook