

מפת"ח • MethodA

## מחסן נתונים DWH

מצגת מנהלים



© כל הזכויות שמורות

[www.methoda.com](http://www.methoda.com)

בחירה באיכות  
**מתודה**

- ❖ מחסן נתונים הוא בסיס נתונים יחסי גדול מאד, המתוכנן לביצוע שאילתות וניתוח נתונים, ולא לעיבוד טרנזקציות
- ❖ מאופיין לרוב בנפחי מידע גדולים של נתונים היסטוריים רב-תקופתיים
- ❖ מחסן נתונים מתבסס על מאגרי מידע אופרטיביים בארגון ומחוצה לו ואינו מעדכן נתונים

❖ הדרך השכיחה להקמת מחסן נתונים היא:

- גזירת הנתונים מהמערכות התפעוליות
- הסבת הנתונים וטיובם
- התאמת הנתונים לסוגי השאילתות ואחזור המידע הנדרשים בארגון

❖ לצורך שליפת הנתונים והצגתם, משולבים במערכת כלים אנליטיים לתחקור והצגת הנתונים, כולל אפשרויות חיפוש מידע מזדמן

# מערכת ענפית - Data Mart

❖ מחסן נתונים המתמקד בתחום עסקי אחד של הארגון:

◦ שכר

◦ כ"א

◦ ייצור

❖ בהקמת מערכת מסוג זה יש לענות על הצרכים הייחודיים של התחום תוך התאמה לצורכי המשתמשים

# מערכת ארגונית - Data Warehouse

- ❖ מחסן נתונים המתמקד במספר תחומים עסקיים
- ❖ לכל תחום יש מאפיינים ייחודיים משלו כאשר חלק מהמאפיינים משותפים למספר תחומים
- ❖ המערכת צריכה לענות על הצרכים הייחודיים של כל ענף
- ❖ המערכת חייבת להתייחס לאינטגרציה של סוגי המידע השונים לכלל מערכת אחת
- ❖ יש לבצע התאמה לצרכים המגוונים של משתמשים רבים בעלי דרישות שונות

# ארכיטקטורה של מערכות מחסן נתונים - מערכות תפעוליות

❖ המערכות האופרטיביות בארגון:

◦ מערכות ניהול הייצור

◦ מערכות כספים

◦ מערכות ליבה - Core Business

❖ מערכות משיקות, חיצוניות לארגון מהוות גם הן מקורות מידע למחסן הנתונים

❖ ממשק האחזור והניתוח מכיל כלים ויישומים לתחקור

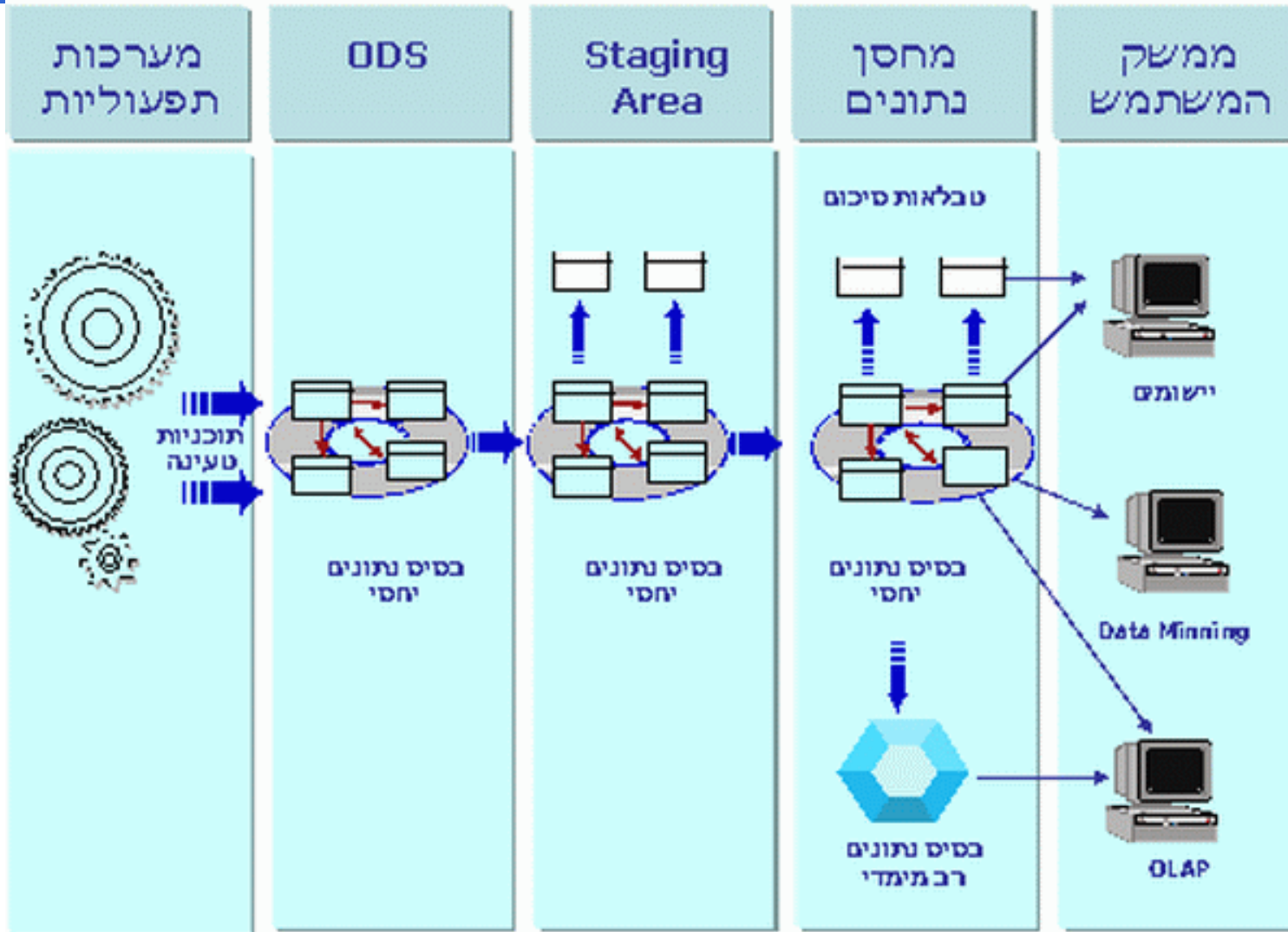
המידע:

- כלי שאילתות
- מחוללי דו"חות
- יישומי Olap (On-Line Analytical Processing)
- כלים לכריית נתונים
- יישומים אופרטיביים משלימים

- ❖ בסיס הנתונים המרכזי של מחסן הנתונים
- ❖ המאגר מכיל את הנתונים שנגזרו מהמערכות התפעוליות, לאחר שטוייבו והותאמו למאגר האינטגרטיבי המרכזי
- ❖ אחסון הנתונים נעשה לרוב במסד נתונים רלציוני, במבנה סכימת כוכב וכולל טבלאות סיכום לשיפור ביצועים
- ❖ יתכן גם קישור לבסיס נתונים רב ממדי ליישומי Olap



# ארכיטקטורת המערכות



## Extraction, Transformation & Load - ETL

❖ שלב מרכזי בהקמת מחסן נתונים. כולל גזירה של הנתונים מהמערכות התפעוליות, ניקויים וטיובם. התאמה של קודים ממערכות ישנות למילון הנתונים המרכזי

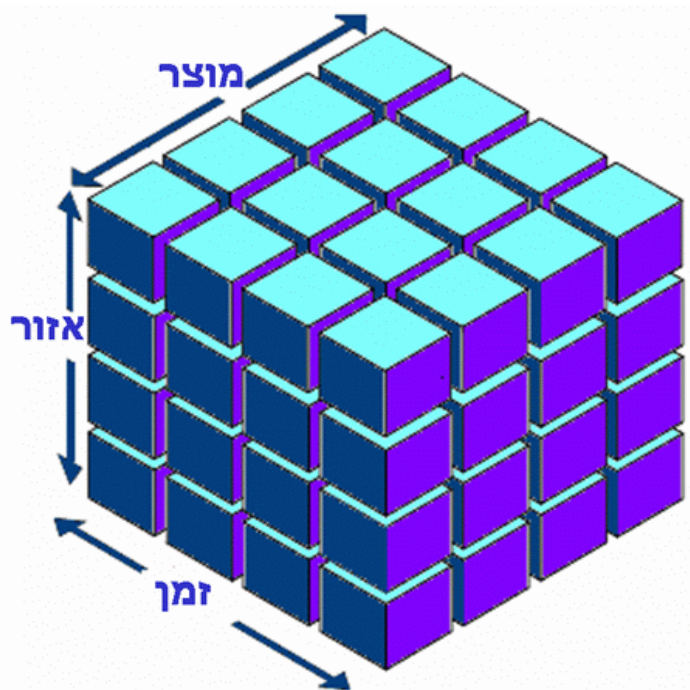
### טבלאות סיכום

❖ אמצעי לשיפור ביצועים של מחסני נתונים גדולים - טבלאות המכילות נתונים מסוכמים לפי ממדים שונים וברמות היררכיה שונות. המחיר - שטחי אחסון גבוהים וזמני טעינה ארוכים

# מונחים מרכזיים (המשך)

## ממדים

❖ מבני נתונים המאפשרים לתחקר את בסיס הנתונים הרב ממדי באספקטים שונים כגון מוצר, אזור וזמן



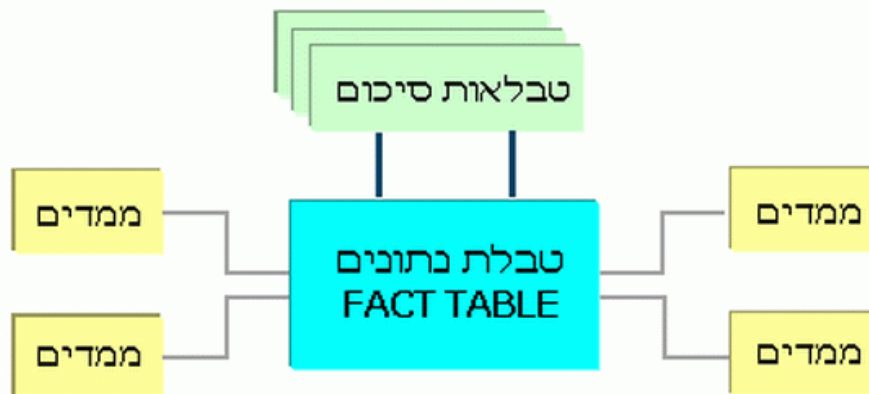
## כריית מידע - Data-Mining

- ❖ תהליך ניתוח אוטומטי המיועד לחשוף תבניות ומגמות. כלי כרייה מכילים יכולות סטטיסטיות ומתמטיות המסייעות בתהליך הניתוח.
- ❖ רוב הכלים מתבססים על עצי החלטה, רשתות נוירונים וסגמנטציה

## סכמת כוכב - Star Schema

❖ פתרון מקובל לעיצוב מחסן נתונים:

- טבלה מרכזית גדולה מאד (Fact Table) בה נמצאות רשומות המידע הבסיסיות
- אוסף של טבלאות אשר מגדירות את הממדים השונים לפיהם ניתן לחתוך את המידע הבסיסי
- טבלאות סיכום המסייעות בביצוע שאילתות מורכבות



# ניהול הפרויקט – מאפיינים

- ❖ פיתוח בסבבים – פיתוח, הקמה, הטמעה
- ❖ אספקה הדרגתית ומהירה של תוצרים
- ❖ הקמה הדרגתית ושילוב בין המערכות בסיום
- ❖ החשיפה לטכנולוגיות המידע החדשות, ולאפשרויות הגלומות במחסן המידע, מביאות את המשתמשים לבקש מידע נוסף ויישומים משלימים

❖ מסגרת הזמן והיקף ההשקעה ארוכים

❖ יש לזהות את המשתמשים הפוטנציאליים ולקיים סקר צרכים מקיף

❖ יש לערוך בדיקה ראשונית של המוצרים הטכנולוגיים הקיימים בשוק

❖ רצוי לבצע בדיקת ישימות ראשונית והקמת אב-טיפוס

❖ המשתמש העיקרי במערכת - Business Analyst, מעורב מאד בכל השלבים

# אפיון

- ❖ מאגר המידע - הגדרה רחבה ככל האפשר של צרכי המידע ודרישות מידע עתידיות
- ❖ אפיון מוקדם של שאילתות ודו"חות
- ❖ עיצוב בסיס הנתונים - מבנה בסיס הנתונים והיקף הנתונים מגבילים את יכולות השליפה
- ❖ הגדרת תהליך Source To Target
- ❖ תכנון המימוש - אפיון וגם הקמה ייעשו בסבבים ובשלבים בהיבט השקעות בתשתיות חומרה וכלי פיתוח ומצד היישום



# בקשה להצעות

- ❖ בחירה נכונה של תשתית טכנולוגית היא גורם מרכזי בהצלחת הפרויקט והבטחת שרידותו
- ❖ יש להתמקד בבסיס הנתונים ובכלי האחזור, השאילתות ומחוללי הדו"חות
- ❖ נושא הסבבים והשקעה הנדרשת באיסוף הנתונים וטיובם חייבים להיות מנוסחים בצורה ברורה במפרט ומוערכים נכונה בין כל ההצעות

## עיצוב ובניה

- ❖ איסוף, טיוב וטעינת הנתונים למאגר המידע כ- 60% מכלל עלויות הקמת המערכת היישומית
- ❖ עיצוב הממשק למשתמש, בניית השאילתות והדו"חות והתאמה של כלי האחזור למודלים הנדרשים של הצגת הנתונים וניתוחם.
- ❖ מאמץ ארגוני להסדרה של העברת הנתונים מהמערכות התפעוליות למחסן הנתונים, וקבלת כל האישורים הנדרשים לכך

# בדיקות מערכת

- ❖ היות שהנתונים משמשים לקבלת החלטות עסקיות, יש להשקיע משאבים רבים בבדיקות
- ❖ קושי באימות הנתונים - המידע במערכת אינו קיים כלל, או קיים בתצורה חלקית במערכות שונות או בדו"חות נייר
- ❖ בדיקת מעבר תקין של הנתונים מהמערכות התפעוליות למאגר המידע ומשם למבנה הנתונים הרב ממדי
- ❖ בדיקה של כלי הקצה (מחוללי דוחות או אפליקציות ייעודיות) לבדיקת תקינות הפלטים

# התקנה והרצה

❖ פיתוח בסבבים גורם למספר "התקנות והרצות"

❖ דגש על הטמעת המערכת אצל המשתמשים

## תפעול ותחזוקה

❖ תהליכים תקופתיים של גזירת הנתונים מהמערכות

התפעוליות, טיובם והבאתם למערכת

## ❖ תיק מערכת/מסמך פרויקט

- תוצר התיעוד המרכזי של המערכת אשר "מתגלגל" לאורך מחזור החיים.
- מורכב מעץ מערכת האוניברסלי של מפת"ח ותוספות ודגשים מיוחדים למערכות מחסן נתונים

## ❖ תיק נושא

- תיק פרויקט לתחום מסוים המהווה חלק ממחסן הנתונים ומקושר לתיק המערכת.

# אבטחת איכות

- ❖ איכות מסמכי הפרויקט בכל שלב במחזור החיים
- ❖ תכנית המימוש של הפרויקט
- ❖ תוצרי אבטחת איכות: נהלים, ניהול סיכונים

# מ תודה

