Indexes

The neglected performance all-rounder

© 2013 by Markus Winand

About Markus Winand

Tuning <u>developers</u> to SQL performance.

Training & co: <u>http://winand.at/</u>

Geeky homepage/blog:
<u>http://use-the-index-luke.com/</u>

Book: "SQL Performance Explained"
(English and German)

New

Translated by:
 Guillaume Lelarge
 (of Dalibo)

SQL:
AU CŒUR DES
PERFORMANCES

ÉDITION FRANÇAISE

COUVRE TOUTES LES BASES DE DONNÉES MAJEURES

TOUT CE QUE LES DÉVELOPPEURS DOIVENT SAVOIR À PROPOS
DES PERFORMANCES EN SQL

MARKUS WINAND

The Problem

Improper Index Use

The Problem: Improper Index Use

"A very common cause of performance problems is <u>lack of proper indexes</u> or the use of <u>queries that are not using existing indexes</u>."

—Buda Consulting

The Problem: Improper Index Use

"A very common cause of performance problems is lack of proper indexes or the use of queries that are not using existing indexes."

—Buda Consulting

Quantifying the Problem

Percona White Paper:

Reasons of performance problems that caused <u>production downtime</u>:

38% bad SQL

15% schema and indexing

Quantifying the Problem

Survey by sqlskills.com:

Root causes of the last few SQL Server performance problems:

27% T-SQL

19% Poor indexing

Quantifying the Problem

Craig S. Mullins (strategist and researcher):

„As much as <u>75%</u> of poor relational performance is caused by <u>"bad" SQL and application code</u>.”

Noel Yuhanna (Forrester Research):

„The <u>key difficulties</u> surrounding performance continue to be <u>poorly written SQL</u> statements, improper DBMS configuration and a lack of clear understanding of how to tune databases to solve performance issues.”

Quantifying the Problem

My observation:

~<u>50%</u> of SQL performance problems
are caused by improper index use

The Root Cause

Admins are Indexing

The Root Cause: DBAs are Indexing

How did we use
indexes before SQL?

The Root Cause: DBAs are Indexing

Index use was intrinsically tied to the queries.

The Root Cause: DBAs are Indexing

Example: dBase

Developers had to...

...use indexes explicitly when searching:

```
set index to last_name
find Winand
```

...take care of index maintenance:
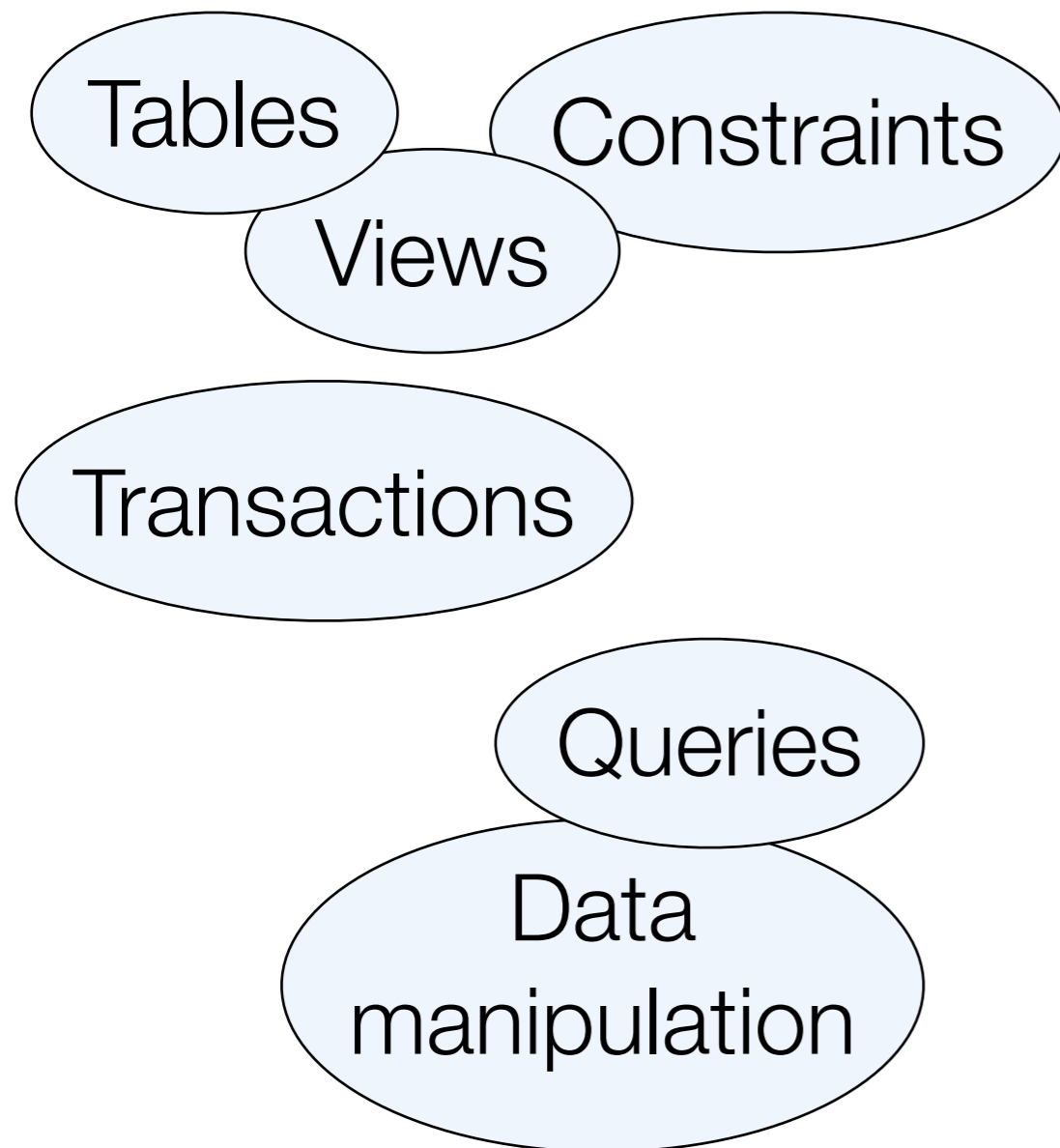
```
set index to last_name, idx2
append
```

The Root Cause: DBAs are Indexing

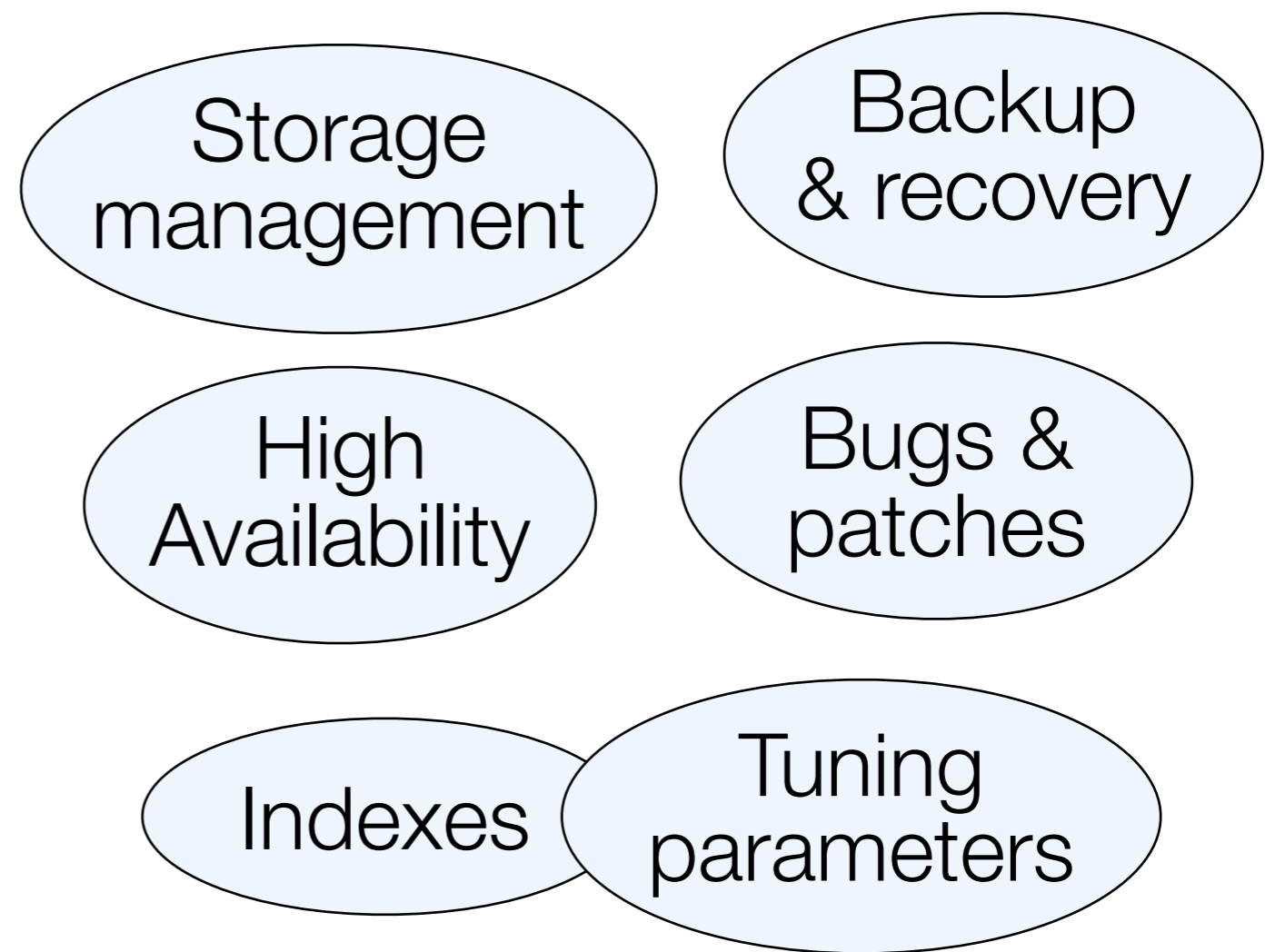SQL is an abstraction that only defines the logical view.

The actual <u>SQL implementation</u> takes care of everything else.

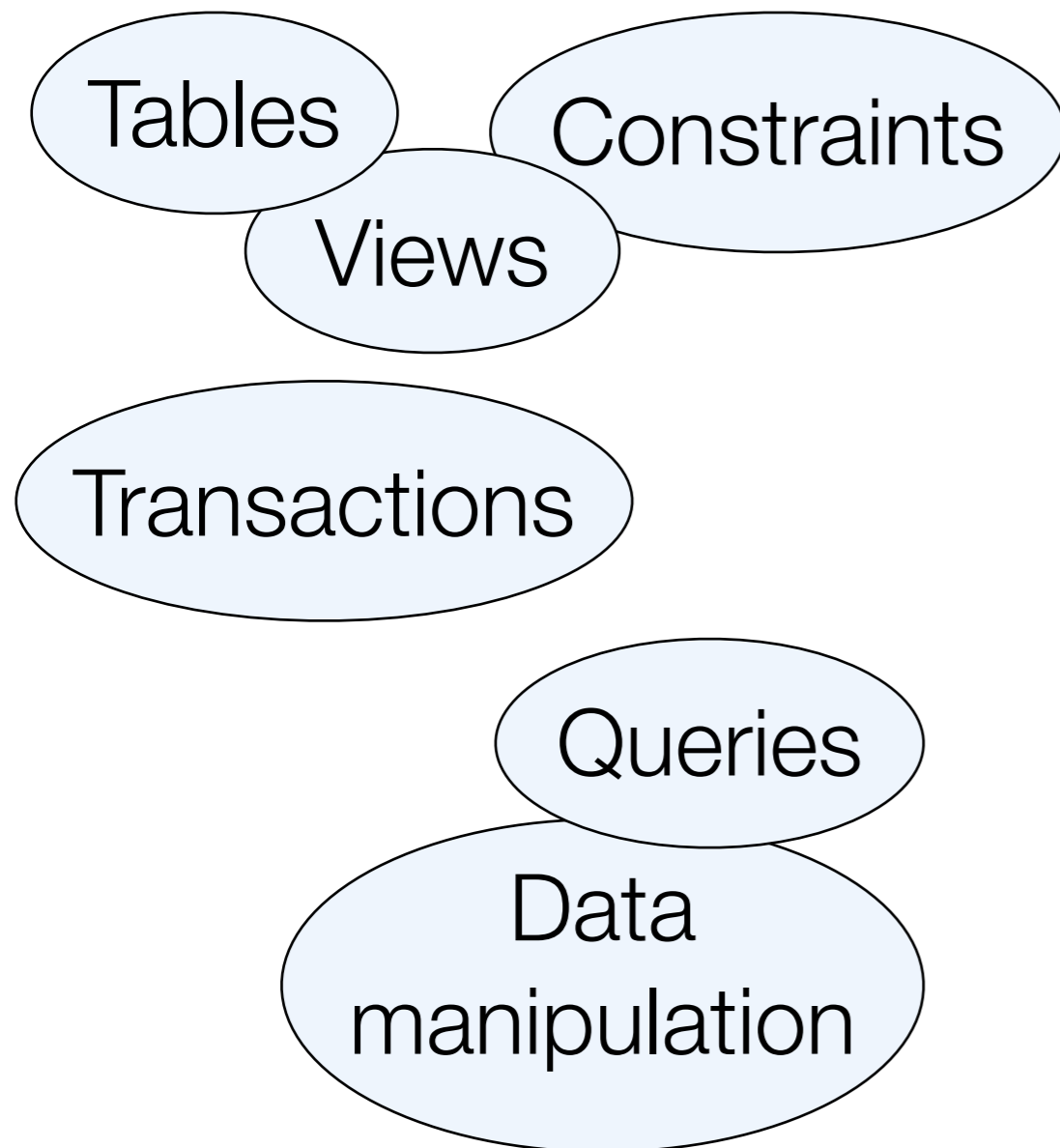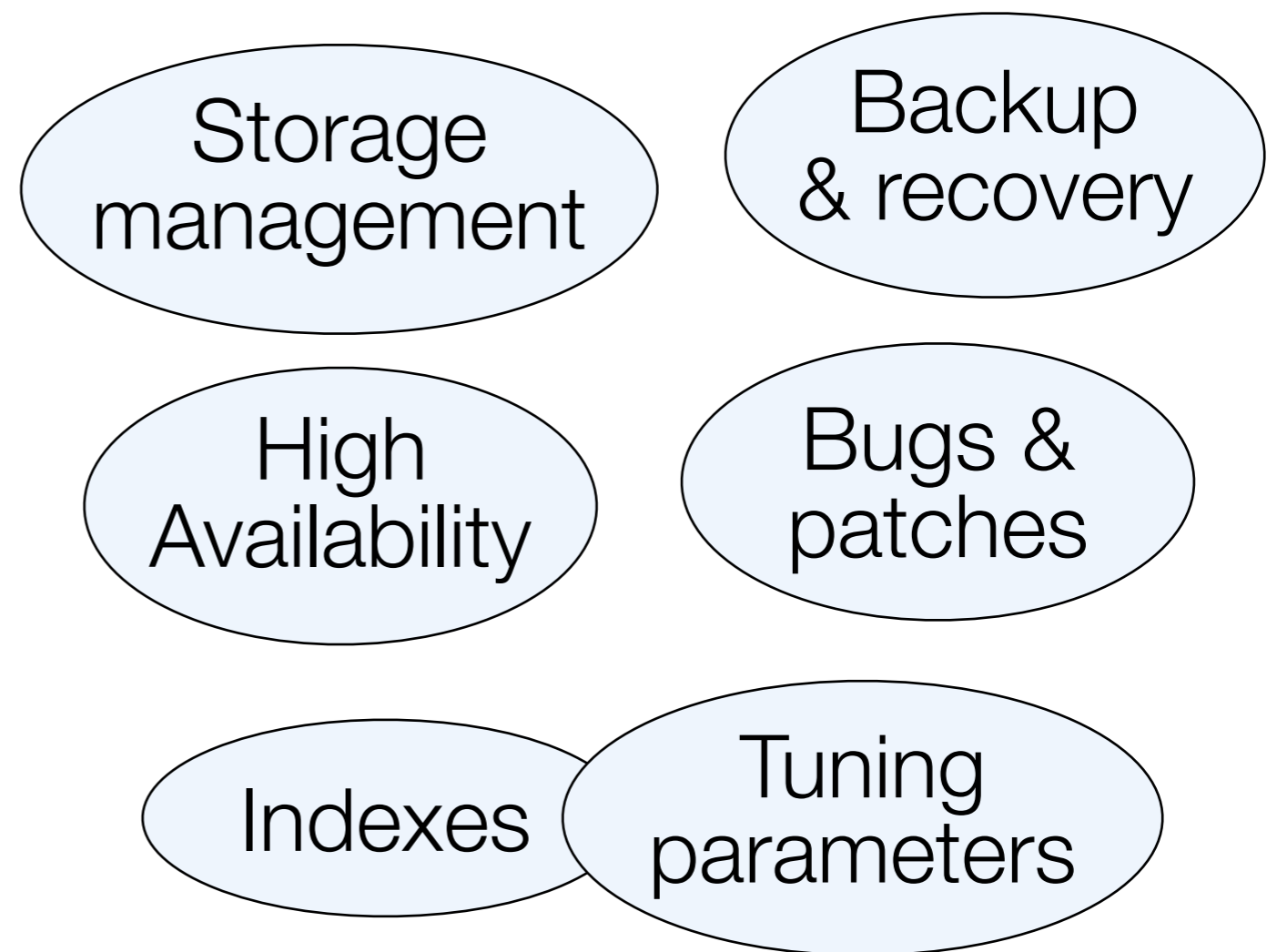# The Root Cause: DBAs are Indexing

SQL (language)

Tables

Constraints

Views

Transactions

Queries

Data manipulation

SQL databases (software)

Storage management

Backup & recovery

High Availability

Bugs & patches

Indexes

Tuning parameters

The Root Cause: DBAs are Indexing

Developers

Tables

Constraints

Views

Transactions

Queries

Data manipulation

SQL databases (software)

Storage management

Backup & recovery

High Availability

Bugs & patches

Indexes

Tuning parameters

The Root Cause: DBAs are Indexing

Developers

Tables

Constraints

Views

Transactions

Queries

Data manipulation

Administrators

Storage management

Backup & recovery

High Availability

Bugs & patches

Indexes

Tuning parameters
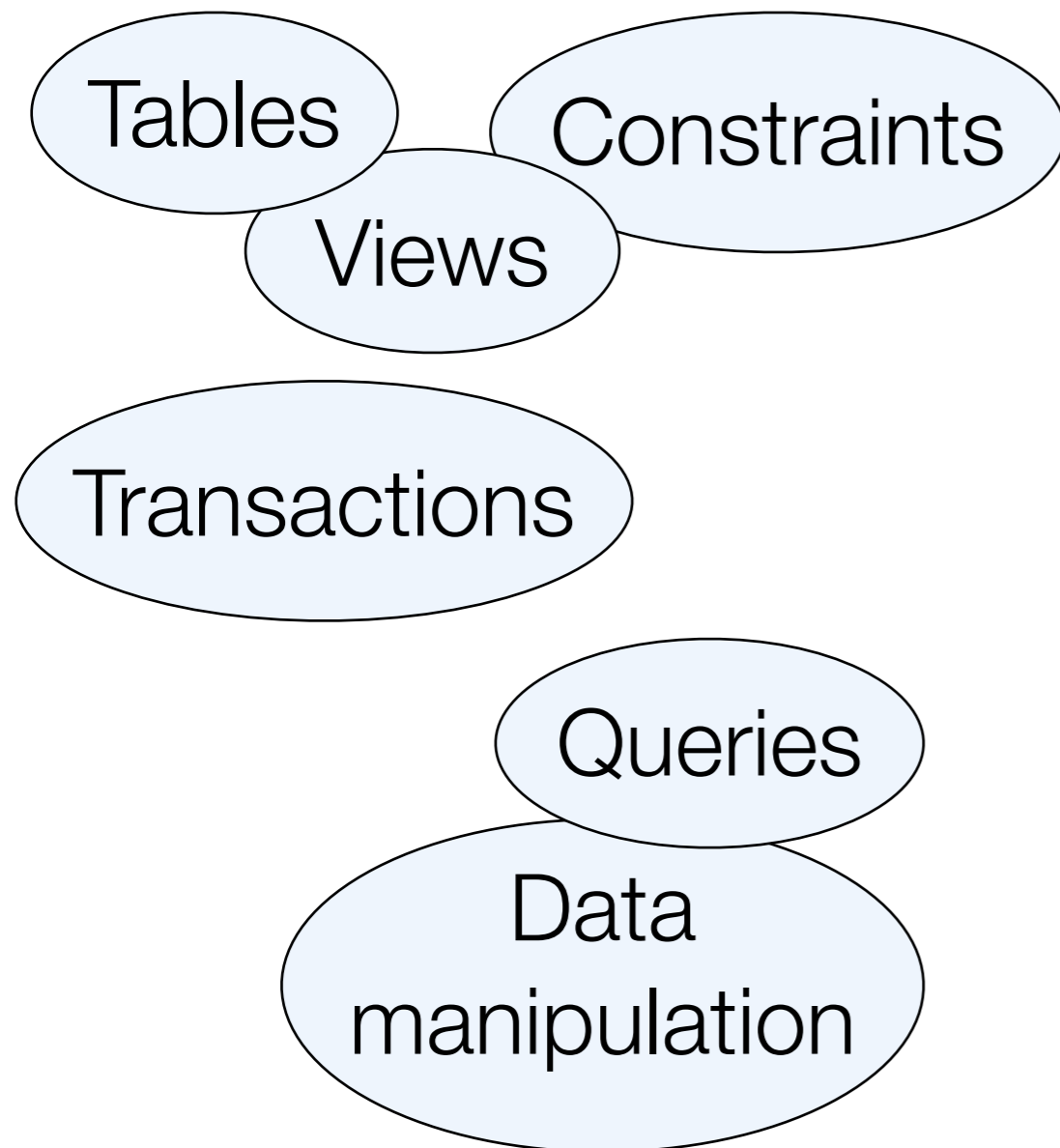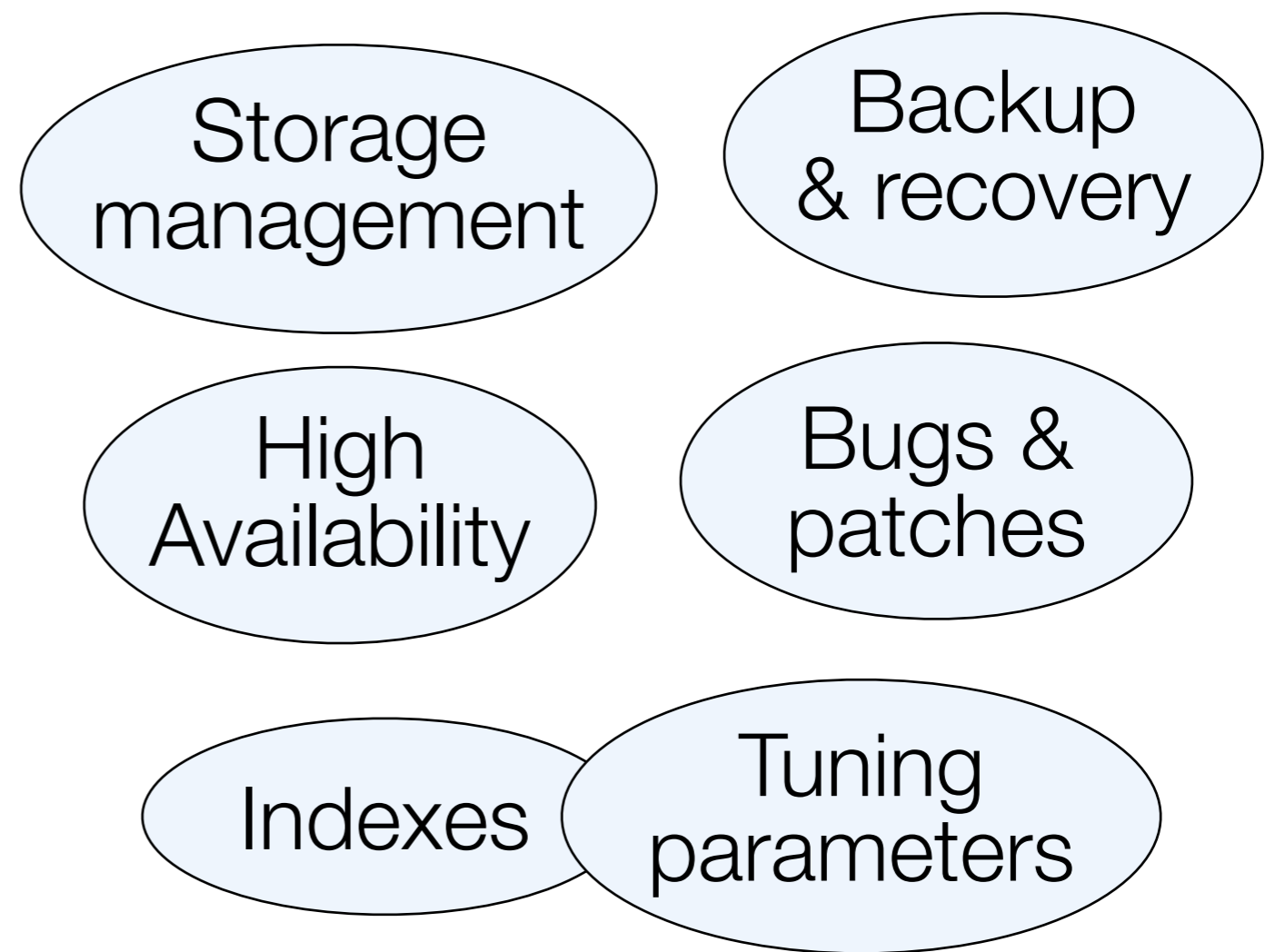
The Root Cause: DBAs are Indexing

Indexing is considered a system tuning task that belongs to the administrators responsibilities.

The Root Cause: DBAs are Indexing

A misconception that causes new problems:

The Root Cause: DBAs are Indexing

A misconception that causes new problems:

DBAs don't know
the queries

Have to "reverse engineer"
the app to get the queries.

It is time consuming and
almost always incomplete.

The Root Cause: DBAs are Indexing

A misconception that causes new problems:

DBAs don't know the queries	DBAs can't change the queries
Have to "reverse engineer" the app to get the queries.	Can make the index match the query.
It is time consuming and almost always incomplete.	Can't make the query match the index!

The Solution

Indexing is a
Development Task

© 2013 by Markus Winand

The Solution: It's a Dev Task

Developers

Tables

Constraints

Views

Transactions

Queries

Indexes

Data manipulation

Must match!

Administrators

Storage management

Backup & recovery

High Availability

Bugs & patches

Tuning parameters

Another Problem: It's not Taught

Indexes are not part of the SQL literature because they are not part of the SQL language (standard).

11 SQL books analyzed: less than **1%** of the pages are about indexes (70 out of 7330 pages).

Examples:

Oracle SQL by Example: **2.0%** (19/960)

Beginning DBs with PostgreSQL: **0.8%** (5/664)

Learning SQL: **3.3%** (11/336 — highest rate in class)

Another Problem: It's not Taught

Proper index usage is <u>sometimes</u> covered in tuning books but is <u>always</u> buried between hundreds of pages of HW, OS and DB parameterization.

14 database administration books analyzed: **5.1%** of the pages are about indexes (307 of 6069 pages).

Examples:

Oracle Performance Survival Guide: **5.2%** (38/730)

High Performance MySQL: **8%** (55/684)

PostgreSQL 9 High Performance: **5.8%** (27/468)

Another Problem: It's not Taught

Consequence:
People don't know how to
use indexes properly.

Results of the 3-minute online quiz:
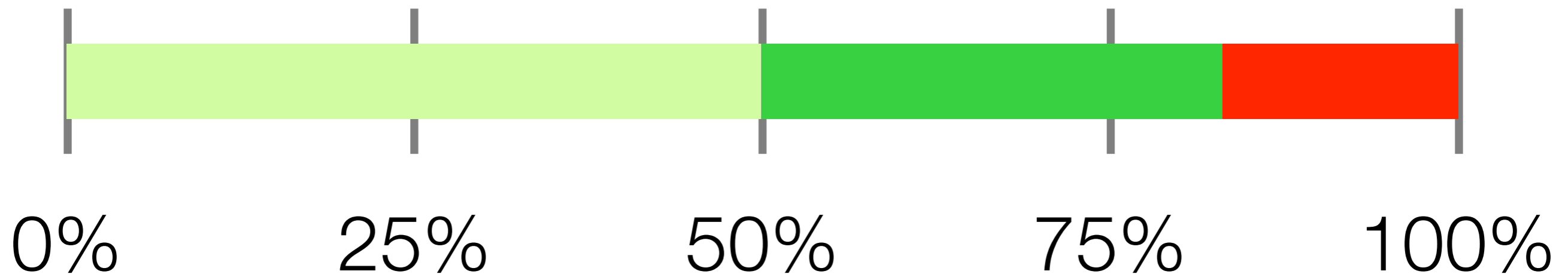http://use-the-index-luke.com/3-minute-test
5 questions: each about a specific
index usage pattern.
Non-representative!

3-Minute Quiz: Indexing Skills
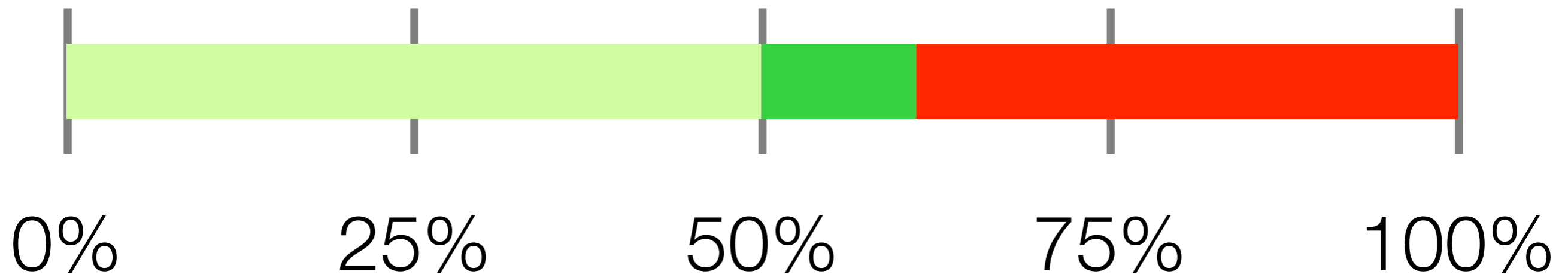
Q1: Good or Bad?   *(Function use)*

```
CREATE INDEX tbl_idx ON tbl (date_column);
SELECT text, date_column
   FROM tbl
 WHERE TO_CHAR(date_column, 'YYYY') = '2012';
```

0%   25%   50%   75%   100%

3-Minute Quiz: Indexing Skills

Q2: Good or Bad?   *(Indexed Top-N, no IOS)*

```
CREATE INDEX tbl_idx ON tbl (a, date_col);
SELECT id, a, date_col
  FROM tbl
 WHERE a = $1
 ORDER BY date_col DESC
 LIMIT 1;
```
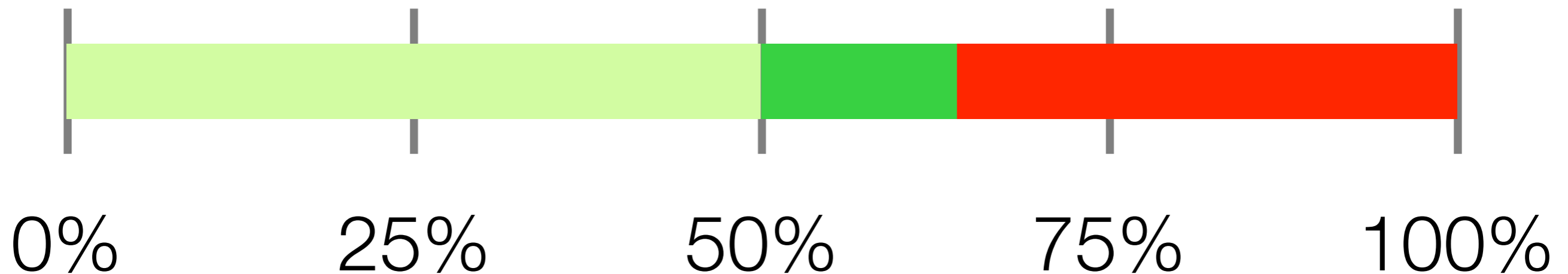
0%   25%   50%   75%   100%

3-Minute Quiz: Indexing Skills

Q3: Good or Bad? *(Column order)*

```
CREATE INDEX tbl_idx ON tbl (a, b);

SELECT id, a, b FROM tbl
 WHERE a = $1 AND b = $2;

SELECT id, a, b FROM tbl
 WHERE b = $1;
```
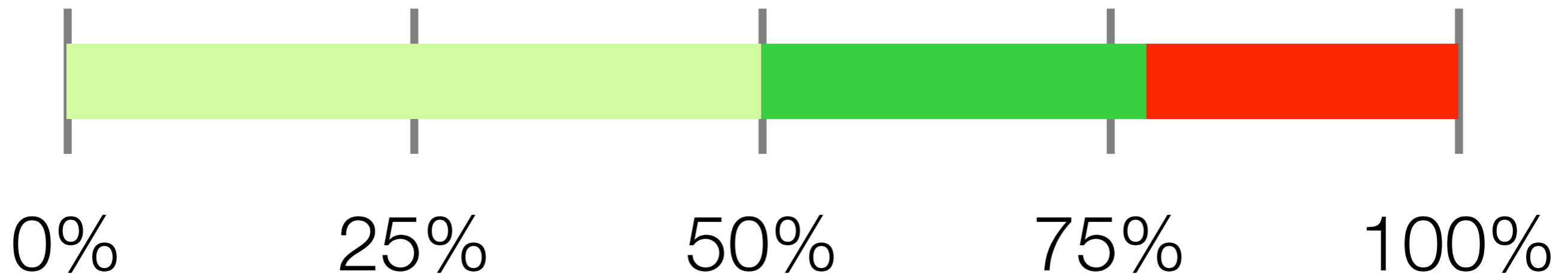
0%   25%   50%   75%   100%

3-Minute Quiz: Indexing Skills

Q4: Good or Bad? *(Indexing LIKE)*

```
CREATE INDEX tbl_idx
    ON tbl (text varchar_pattern_ops);

SELECT id, text
  FROM tbl
 WHERE text LIKE '%TERM%';
```
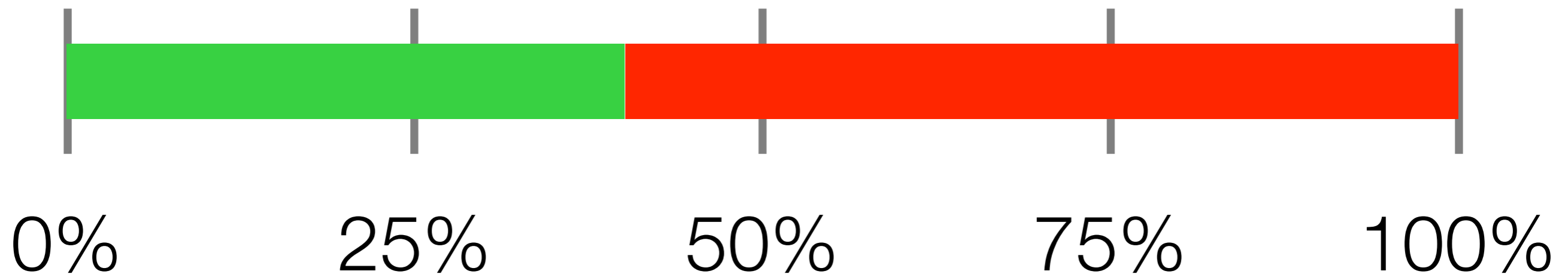
0%   25%   50%   75%   100%

3-Minute Quiz: Indexing Skills

Q5: Good or Bad?   *(equality vs. ranges)*

```
CREATE INDEX tbl_idx
        ON tbl (date_col, state);

SELECT id, date_col, state FROM tbl
  WHERE date_col >=
        CURRENT_DATE - INTERVAL '5' YEAR
    AND state = 'X';
```

0%   25%   50%   75%   100%

The Neglected All-Rounder

Everybody knows indexing is important for performance, still nobody takes the time to learn and apply it properly.

The Neglected All-Rounder

Index details are hardly known.
➡ "Details" like column-order or equality vs. range conditions must be <u>learned and understood</u>.

Only one index capability is used: finding data quickly
➡ Indexes have <u>three capabilities (powers)</u>: finding data, clustering data, and sorting data.

Indexing is done from single query perspective.
➡ Should be done from application perspective (considering all queries). It's a <u>design</u> task!

The Neglected All-Rounder

Are you just adding indexes

or

are you designing indexes?

http://Use-The-Index-Luke.com/