1 Content

2 Introduction

dfdgfgfhg hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh loremdfdsfdfre

Definition

Dummy data is mock data generated at random as a substitute for live data in testing environments. In other words, dummy data acts as a placeholder for live data, the latter of which testers only introduce once it's determined that the trail program does not have any unintended, negative impact on the underlying data.

For example, a company implementing a new accounting system uses dummy data to ensure its bookings are stored correctly before inputting live accounts.

You can think of dummy data like a car crash dummy — you hope nothing bad happens to it while testing, but if something bad does happen, you can fix the problem with no negative impact on your valuable data assets.

Make Sure You Know Your Dummy Data

A classic slip-up for new analysts is not knowing dummy data from real data, which be confusing (and embarrassing) when you're working fast on a project with others. That's because in some cases dummy data can "look" real, so make sure to ensure everyone involved is aware when dummy data is active.

Don't forget, you can access the 67 data skill and concepts checklist for free to ensure you've got your bases covered.

Examples of Dummy Data

Analysts use dummy data in two primary contexts (we'll look at a third later): testing new programs and testing modifications on existing program. Let's look at an example of each below.

Imagine you own a big e-commerce website that sells watches in bulk. Your company is called Batch Watch — it's your baby. One of the biggest assets the company has is the data it collects on vendors, its products, and customers.

When you started the company, you didn't have much cash to spend on an expensive database program, so you took the first inexpensive option that came your way. Unfortunately, this means your data is not well-modeled, and you're not really sure howyou didn't have much cash to spend on an expensive database program, so you took the first inexpensive option that came your way. Unfortunately, this means your data is not well-modeled, and you're not really sure how that initial data program's technical skeleton is structured.