

Katusome IDEA practice datasets: data dictionary

All data is fictional and was generated for training. Company: Kampala Trading Co. Ltd (fictional). Any resemblance to real people or businesses is coincidental. Reference "as at" date: 31 December 2025.

File	Rows	Purpose
hr_master.csv	620	HR master list of employees of Kampala Trading Co. Ltd (fictional). Source of truth for who works there and when they left.
system_users.csv	670	User accounts extracted from the ERP. Compare with HR for leavers, orphans, duplicates, password and dormancy tests.
access_requests.csv	640	Approved access request forms. Compare approval date with the date access was granted.
revocation_requests.csv	120	Leaver access revocation tickets: HR termination date, request date and the date the account was actually disabled.
roles.csv	19	List of ERP roles.
user_roles.csv	742	Which user holds which role (a user can hold several).
sod_conflicts.csv	11	Segregation of duties conflict matrix: role pairs that should not be held by one person.
approver_limits.csv	28	Payment approval limit in UGX per approver.
transactions.csv	5000	Payment and journal vouchers for 2025 with maker, checker and timestamps.
vendors.csv	150	Vendor master file.
vendor_changes.csv	296	Change log of vendor master data (bank details, phone, etc.).
purchase_orders.csv	800	Purchase orders.
grn.csv	775	Goods received notes.
invoices.csv	825	Supplier invoices captured for payment.
payroll.csv	516	December 2025 payroll run. Simplified practice formulas, see notes below.
change_log.csv	900	IT change management register for 2025.
backup_log.csv	1797	Daily backup job log, five systems, 2025.
incident_tickets.csv	1540	Service desk incident tickets for 2025.
admin_audit_log.csv	3610	Administrator activity log extract for 2025 (log_id is sequential).
patch_status.csv	800	Asset patch and vulnerability status as at 31 Dec 2025.
gl_vs_subledger.csv	60	Month-end GL control account balance against the subledger total, 2025.
control_totals.csv	21	Record counts and control totals supplied by the (fictional) source systems. Use to prove your import is complete.

hr_master.csv

Field	Description
emp_id	Employee ID (links to hr_master)
name	Full name (fictional)
department	Descriptive field name
job_title	Descriptive field name
hire_date	Date of hire
termination_date	Date employment ended (blank if still employed)
status	Status
manager_id	Descriptive field name

bank_name	Descriptive field name
bank_account	Bank account number (fictional)

system_users.csv

Reference date for age calculations is 31 Dec 2025 (the audit "as at" date). emp_id is blank for generic and service accounts; owner_emp_id shows the named owner where one exists.

Field	Description
user_id	User ID
username	ERP login name
emp_id	Employee ID (links to hr_master)
owner_emp_id	Descriptive field name
full_name	Descriptive field name
account_type	Descriptive field name
account_status	Active or Disabled
role_id	Role ID (links to roles)
privilege_level	Standard, Elevated or Admin
created_date	Descriptive field name
created_by	Username that created the record
password_last_changed	Date the password was last changed (blank = never)
password_never_expires	Y if the password is set not to expire
last_login	Date of last successful login (blank = never)
account_disabled_date	Date the account was disabled

access_requests.csv

Field	Description
request_id	Descriptive field name
emp_id	Employee ID (links to hr_master)
username	ERP login name
role_requested	Descriptive field name
request_date	Descriptive field name
approver_id	Descriptive field name
approval_date	Descriptive field name
access_granted_date	Descriptive field name
status	Status

revocation_requests.csv

Field	Description
revocation_id	Descriptive field name
emp_id	Employee ID (links to hr_master)
username	ERP login name
termination_date	Date employment ended (blank if still employed)
revocation_request_date	Descriptive field name

actual_disable_date	Descriptive field name
status	Status

roles.csv

Field	Description
role_id	Role ID (links to roles)
role_name	Descriptive field name
description	Descriptive field name

user_roles.csv

Field	Description
user_id	User ID
username	ERP login name
role_id	Role ID (links to roles)

sod_conflicts.csv

Field	Description
role_a	Descriptive field name
role_b	Descriptive field name
conflict_description	Descriptive field name
risk	Descriptive field name

approver_limits.csv

Field	Description
username	ERP login name
approval_limit	Maximum UGX the approver may approve

transactions.csv

Amounts in UGX. Approval limit per approver is in approver_limits.csv. Weekend/after-hours means Saturday/Sunday or posting before 06:00 or from 20:00.

Field	Description
voucher_no	Voucher number, sequential
voucher_type	Descriptive field name
vendor_id	Vendor ID (links to vendors)
description	Descriptive field name
amount	Amount in UGX
currency	Descriptive field name
created_by	Username that created the record
created_datetime	When the voucher was created
approved_by	Username/ID that approved the record
approved_datetime	When the voucher was approved

posting_datetime	When the voucher was posted to the ledger
status	Status

vendors.csv

Field	Description
vendor_id	Vendor ID (links to vendors)
vendor_name	Descriptive field name
tin	Descriptive field name
bank_name	Descriptive field name
bank_account	Bank account number (fictional)
phone	Descriptive field name
created_by	Username that created the record
created_date	Descriptive field name
status	Status

vendor_changes.csv

Field	Description
change_id	Change reference
vendor_id	Vendor ID (links to vendors)
field_changed	Descriptive field name
old_value	Descriptive field name
new_value	Descriptive field name
changed_by	Descriptive field name
change_datetime	Descriptive field name
approved_by	Username/ID that approved the record
approval_datetime	Descriptive field name

purchase_orders.csv

Field	Description
po_no	Purchase order number
vendor_id	Vendor ID (links to vendors)
po_date	Descriptive field name
item	Descriptive field name
qty	Descriptive field name
unit_price	Descriptive field name
po_amount	Descriptive field name
created_by	Username that created the record
approved_by	Username/ID that approved the record
po_approval_date	Descriptive field name

grn.csv

Field	Description
grn_no	Descriptive field name
po_no	Purchase order number
grn_date	Descriptive field name
qty_received	Descriptive field name
received_by	Descriptive field name

invoices.csv

Field	Description
inv_id	Descriptive field name
invoice_no	Descriptive field name
po_no	Purchase order number
vendor_id	Vendor ID (links to vendors)
invoice_date	Descriptive field name
qty_invoiced	Descriptive field name
unit_price	Descriptive field name
net_amount	Descriptive field name
vat_amount	Descriptive field name
total_amount	Descriptive field name
captured_by	Descriptive field name

payroll.csv

Practice formulas: gross = basic + allowances; NSSF employee 5% of gross; PAYE bands (monthly, simplified): 0 up to 235,000; 10% of excess over 235,000 up to 335,000; 10,000 + 20% of excess over 335,000 up to 410,000; 25,000 + 30% of excess over 410,000 up to 10,000,000; plus 40% above 10,000,000 (simplified for practice, check URA tables for real work); net = gross - PAYE - NSSF employee.

Field	Description
payroll_month	Descriptive field name
emp_id	Employee ID (links to hr_master)
name	Full name (fictional)
basic_salary	Descriptive field name
allowances	Descriptive field name
gross_pay	Descriptive field name
paye	Descriptive field name
nssf_employee	Descriptive field name
nssf_employer	Descriptive field name
net_pay	Descriptive field name
bank_account	Bank account number (fictional)
pay_date	Descriptive field name

change_log.csv

Field	Description
-------	-------------

change_id	Change reference
system	Descriptive field name
description	Descriptive field name
change_type	Descriptive field name
requested_by	Descriptive field name
implemented_by	Descriptive field name
request_date	Descriptive field name
approver	Descriptive field name
approval_date	Descriptive field name
test_date	Descriptive field name
test_result	Pass, Fail or Not tested
deployed_date	Date change went to production
deployed_by	Descriptive field name

backup_log.csv

Field	Description
job_id	Descriptive field name
system	Descriptive field name
backup_date	Descriptive field name
start_time	Descriptive field name
end_time	Descriptive field name
status	Status
size_gb	Descriptive field name
notes	Descriptive field name

incident_tickets.csv

Field	Description
ticket_id	Descriptive field name
category	Descriptive field name
priority	Descriptive field name
sla_hours	Resolution target in hours for the priority
opened_datetime	Ticket opened
resolved_datetime	Ticket resolved (blank = not resolved)
closed_datetime	Ticket closed
status	Status
assigned_to	Descriptive field name
resolution_notes	Descriptive field name

admin_audit_log.csv

Field	Description
log_id	Sequential log number

event_datetime	Time of the event
username	ERP login name
action	Action performed
target	Descriptive field name
source_ip	Descriptive field name

patch_status.csv

Field	Description
asset_id	Descriptive field name
hostname	Descriptive field name
asset_type	Descriptive field name
os	Descriptive field name
os_support_status	Descriptive field name
owner_emp_id	Descriptive field name
department	Descriptive field name
criticality	Descriptive field name
last_patch_date	Descriptive field name
last_scan_date	Last vulnerability scan date
pending_critical_patches	Descriptive field name
oldest_pending_released	Release date of the oldest missing critical patch
max_cvss	Descriptive field name

gl_vs_subledger.csv

Field	Description
period	Descriptive field name
area	Descriptive field name
gl_balance	General ledger balance
subledger_balance	Subledger total
difference	GL minus subledger

control_totals.csv

Field	Description
file	Descriptive field name
record_count	Descriptive field name
control_field	Descriptive field name
control_total	Descriptive field name