

Busting the Census count myth

Owen Abbott

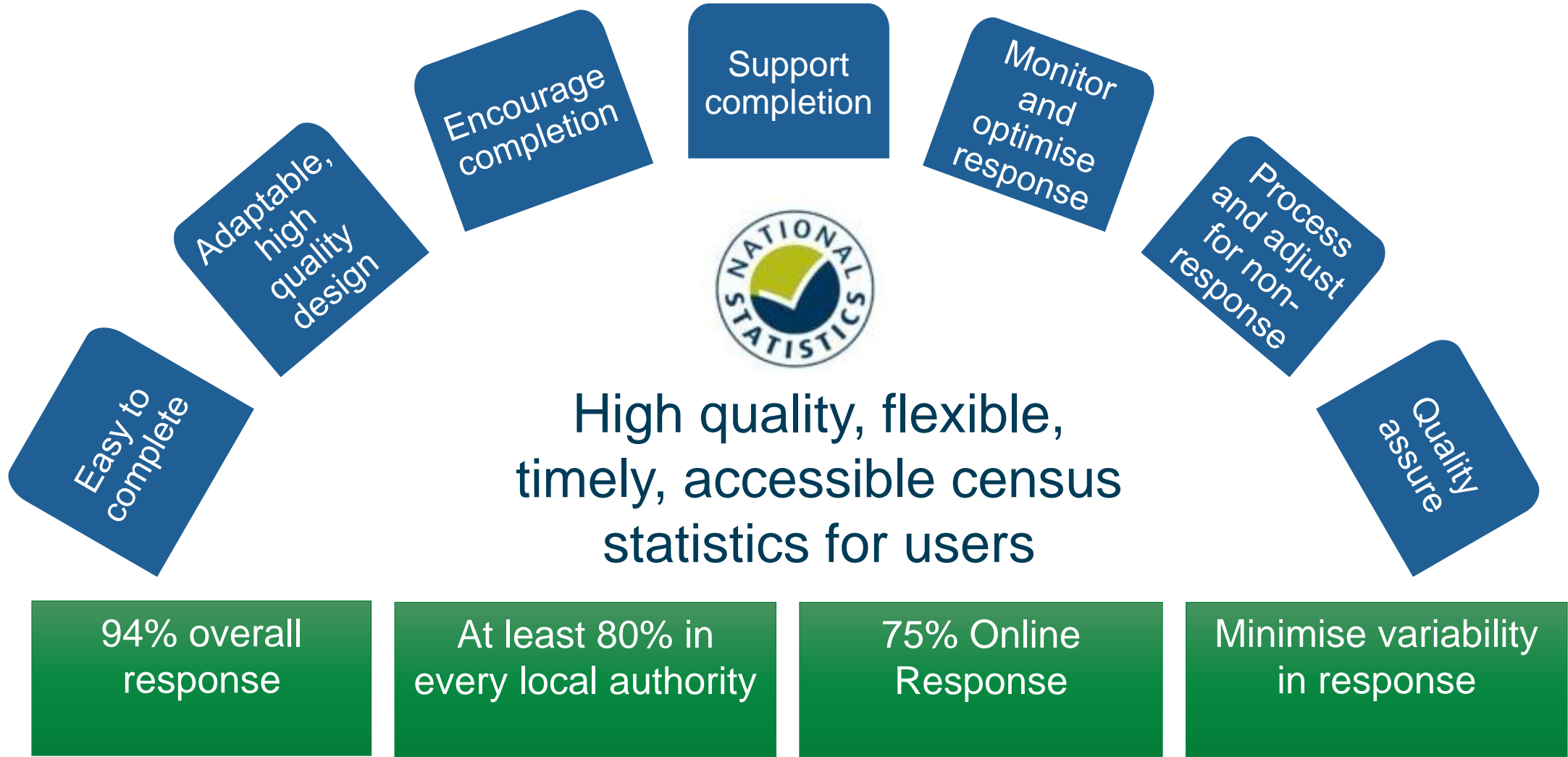
Structure

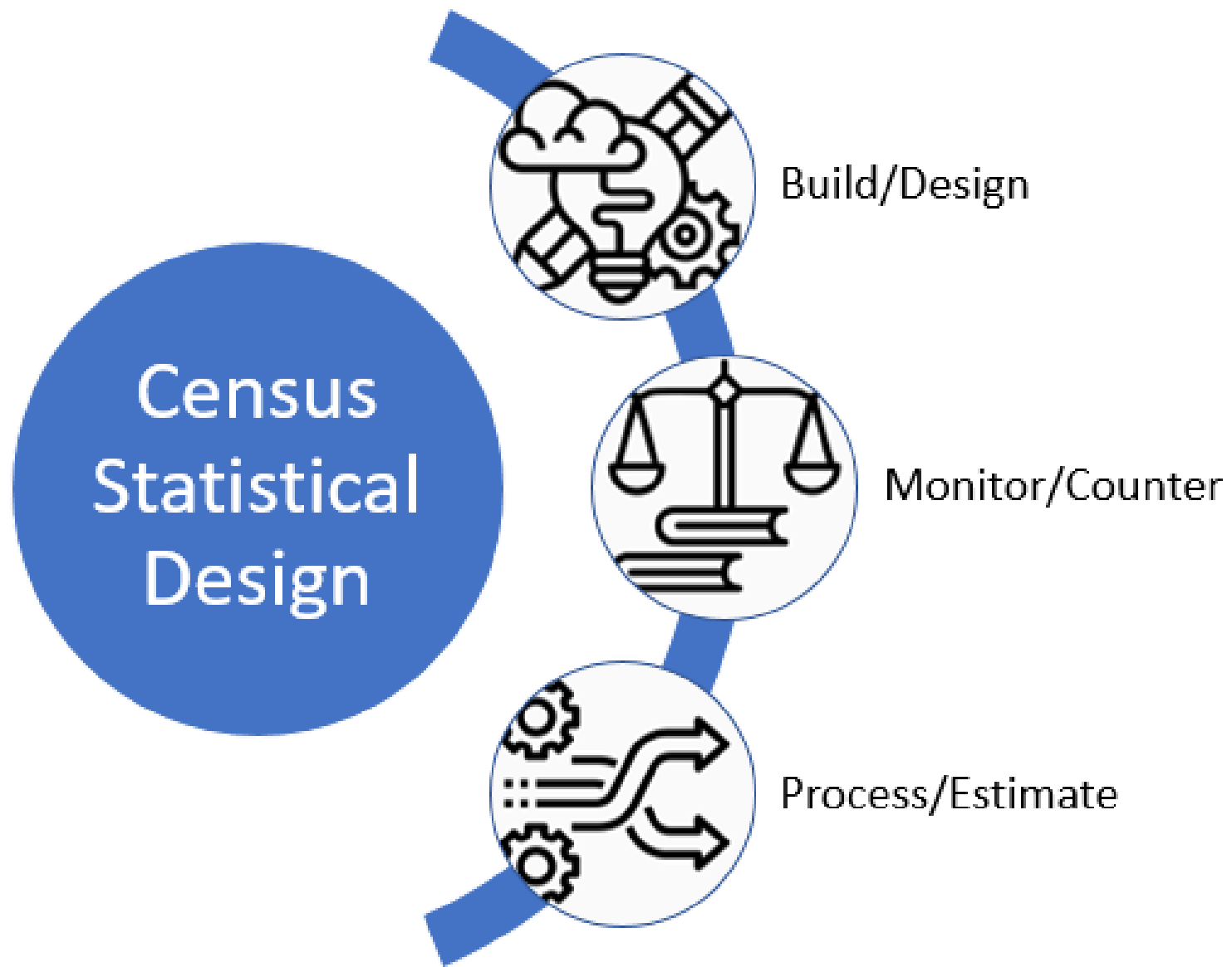
- The 2021 Census in England and Wales
- Statistical processes
- Challenges for 2021 Census
- Quality measurement

census 2021

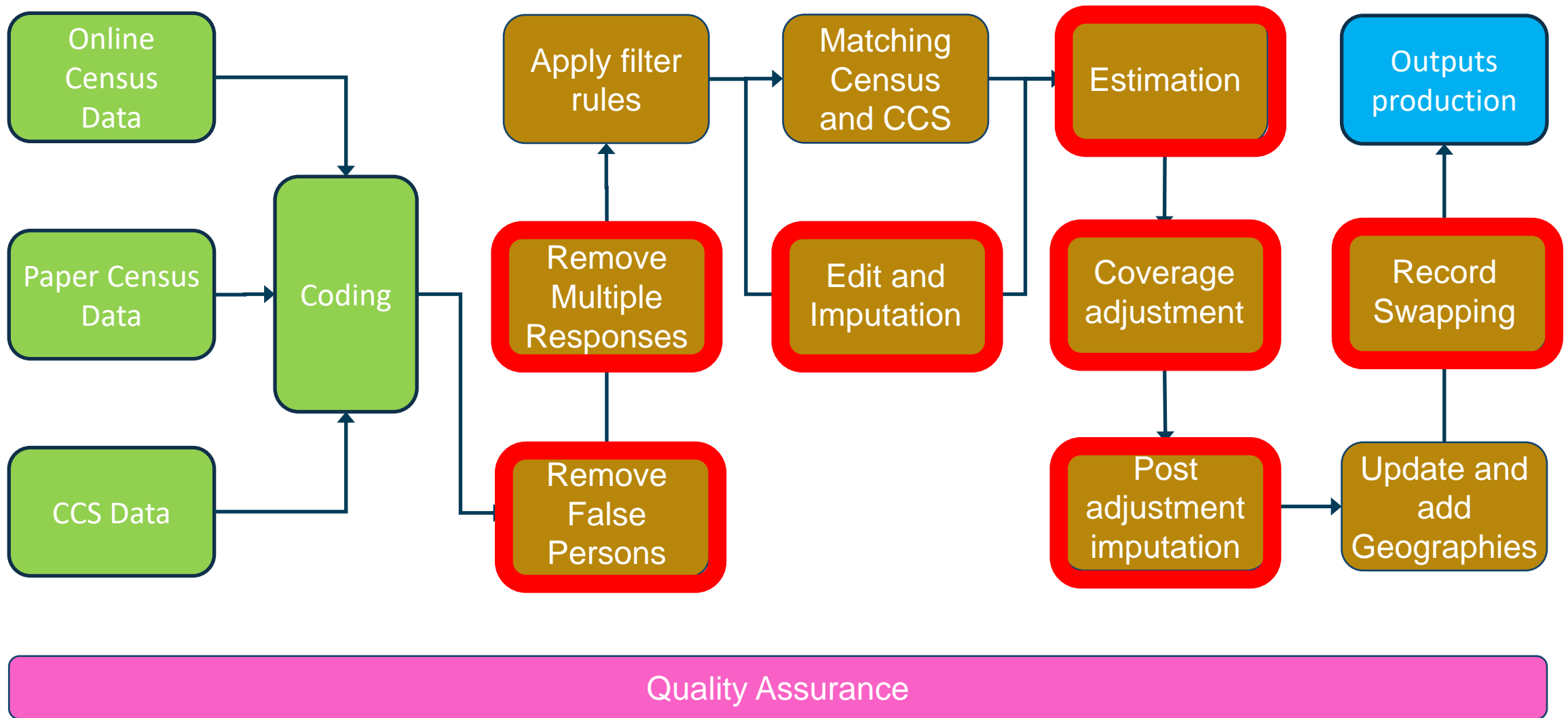
CENSUS SIMPLIFIED







High level processing components and flows



Data Cleaning (1)

- Remove False Persons – where there is insufficient data
- Remove Multiple Responses – where there are duplicates within an address, or wholly duplicated addresses

Data Cleaning (2)

- Edit – Apply multi-tick rules, range checks and edit rules to remove inconsistencies
- Imputation – statistical imputation of valid values where responses are missing (for persons, households, CEs and CCS). Includes imputation of VOA (Rooms).
- Uses CANCEIS donor imputation system

Coverage Estimation

- Estimates under-coverage and over-coverage
- Uses Census linked to CCS
- National statistical models to produce key population estimates by LA (persons, households, small CEs) and for other characteristics (ethnicity, tenure etc)
- Administrative data used to QA the estimates and assist in some bias adjustments

Coverage Adjustment

- Statistical imputation of whole household records so that the database represents the whole population (households, persons).
- Meets the estimation totals (households, persons, by characteristics)
- Adds people into CEs.
- Uses donor imputation methods.

Statistical Disclosure Control

- Methods developed to protect outputs from identifying individuals, by adding error.
- Balancing risk versus utility
- Pre-tabular record swapping combined with tabular Cell-key method

Challenges for 2021

- COVID:
 - Recruitment and uncertainty in the operation
 - Usual residence concept
 - Students
 - Administrative data
- Higher than expected Census online response
- Lower than expected CCS response

Measuring quality (1)

- Statistical processes remove bias but introduce variance
- All census outputs are estimates!
- Important we communicate this to users
- We produce confidence intervals around the key population estimates

Measuring quality (2)

- We are exploring:
- Imputation variance
- CVs from the adjustment process
- A way to combine these into an overall quality measure
- How to disseminate to our users

Summary

Summary

- The 2021 Census in England and Wales is designed to meet user needs and represent the whole population
- Its statistical design has a number of robust methods to produce unbiased estimates
- It is an integrated design with flexibility to react and mitigate
- It gives us the best chance of a successful 2021 Census
- ...And a great platform for the future
- ...And there are no counts