SESSION 11:
DATA COLLECTION WITH THE INTERNET

DEPARTMENT OF STATISTICS MALAYSIA
Regional Workshop On The 2020 World Programme on Population and Housing Censuses: International Standards and Contemporary Technologies
Ankara, Turki
12-15 March 2019

www.dosm.gov.my @StatsMalaysia
1. BACKGROUND
DOSM has introduced self-response for data collection in Census 2010 with the option of 2 modes:

- **Manual**
- **Online**

**DOPU**

**e-Census**

**Census Kit**

- ID for e-Census system
- through official DOSM Portal

### OUTCOMES

- **Targeted response:** 2.8% (28,000 respondents)
- **Actual response:** <1%
  
  *(Census 2010 Post Mortem Report)*

- **Indirect factor of low internet penetration in 2010 (56.3%)**

Malaysia Census Transformation Programme 2020 (MyCTP 2020)

- Designed towards latest modernised framework which aligned with ICT
- The objective is to accommodate high impact outcomes for sustainable delivery within cost effective
1.2 CENSUS 2020

- **CAPI**: Urban: 33%, Rural: 7%
  - CAPI = Computer Assisted Personal Interview
- **e-CENSUS**: Urban: 21%, Rural: 9%
- **PAPI**: Urban: 11%, Rural: 9%
  - PAPI = Paper Assisted Personal Interview
- **CATI**: Urban: 5%, Rural: -
  - CATI = Computer Assisted Telephone Interview
- **DOPU**: Urban: 5%, Rural: -
  - DOPU = Drop Off Pick Up

ONLINE
1.3 ASPIRATION OF CENSUS 2020

- Increased data security: Security/Risk
- Modernisation of existing processes (e.g., Address Verification System [AVS], GIS) and integration of administrative data with OGSA data.
- Modernisation of census fieldwork using mobile apps and tablets.
- Increase self-response rate: e-Census & DOPU

Rapidly changing use of ICT
Availability of administrative data
Increasingly complex Modus operandi
Changing population profile

Monetary implication – increasing cost
Stakeholders’ expectation & dynamic & demanding Users

Changing Environment

1. More efficient & complete Census Management
2. Application of administrative data
3. Use of latest, more dynamic & flexible ICT infrastructure and application system
4. Open Data Initiative & Big Data Analytics
5. More efficient & long-term capital oriented investment
6. More strategic communication strategy

www.dosm.gov.my  Facebook  Twitter  Instagram  @StatsMalaysia
2. BENEFITS
2.1 BENEFITS

NEW STATISTICAL INFORMATION

1. The basis information for national planning and development
   - Population information, household and living quarters to the smallest geographic area (local statistics)

REGISTER BASED STATISTICS (MSAR & MSPR)

2. Complete LQ and population address database
   - The Population and Housing Census can be implemented on a smaller scale in the future

STATISTICS LANDSCAPE TRANSFORMATION

3. Improve the efficiency of statistical delivery systems to be more modern, dynamic and user-friendly
   - Improve the statistical literacy
2.1 BENEFITS

CENSUS PARAMETER SHARING
- Data sharing aspects where most updated data from MSAR & MSPR can be shared with various agencies (aggregate data)
- Cooperation with Other Government Agencies (OGA) in the use of administrative data can be enhanced

REGISTER BASED STATISTICS (MSAR & MSPR)
- Population input for the establishment of key national policy such as National Transformation 2050 (TN50) and population aging

EFFECTIVE COST
- Long term investments
- Continuous ICT investment
- Cost effective for future
2.2 CHALLENGES

Data governance

Publicity & promotion

Normalisation & adoption of procedures in the fieldwork based on technology implemented

Data Security

Data Risk

Modernisation of census data collection:
OSL, FL, eCensus, CAPI, CATI, CT & Portal

Modernisation of field work:
FL, CAPI

Centralised integrated administrative data:
MSAR & MSPR

Increase Self-Response Rate:
OSL, eCensus, CT

Online Portal as one stop center for data dissemination of pre, during & post census

Complex & stabilised ICT infrastructure

Peak response rate in relates with the factor of internet coverage, especially for last minutes response

Online Learning Hub
Malaysia Statistics Ubiquitous Learning (MySUL)
3. DATA COLLECTIONS
3.1 DATA COLLECTIONS IN MyIPCS

5 Modes Enumerations
- 3 Computer Assisted
  • e-Census (Web)
  • CATI (Web)
  • CAPI (Mobile: Offline/Online)
- DOPU & PAPI

Integrated Database
- Data Management done via system
- Geocoded Data
- Address ID Generation

Smart Listing & Smart Map
- Listing via Web/Tablet
- OSL system for public
- Update coordinate at LQ Level
- Improve location using GPS

Smart Search
- Secondary Data Source
- Online Search & Verification

System Notification
- Email

Dual Language
- English
- Malay

Analytics & Infographic
- Advance Analytic Tools
- Interactive Visualization Dashboard

Web Responsive
- Self Adaptive to all screen resolution & smart devices
- User friendly interface

Code Classification
- Algorithm Text Coding
- Checking Data Dictionary

www.dosm.gov.my  @StatsMalaysia
3.2 MODULES IN MyIPCS

ADMINISTRATION & INTEGRATION

CENSUS PROCESSING
5. Process of consolidation, verification, validation, imputation and updating raw census data from Census Operation

CENSUS DATA MANAGEMENT
2. Execute & manage data harmonization and standardize addresses and individual data & establishment of MSAR, MSFR and Frozen Frame. And, generation of unique ID to represent address

CSS : QUALITY CONTROL
6. Online platform to do QC Census Operation and QC Census Processing activities

ANALYTICS & REPOSITORY
9. To perform the analyses that contain descriptive, prescriptive and predictive analysis from data that resides in the available data sources with Analytic Tool

DOSM SMART LISTING & SMART MAP
3. To introduce platform for respondent Online Self Listing (OSL) and Field Listing (FL) for data collection process. And, to cater Geographical Information System (GIS) requirements for spatial data updating and viewing

CSS : CENSUS TRACKING
7. Platform to track, monitor and manage Document, Asset, Inventory, Personnel, Premise and Vendor across different phases of the census

DISSEMINATION & DATA COMM.
10. Visualisation application to produce interactive dashboards, story-telling maps, charts, reports and infographics for data statistic dissemination

CENSUS OPERATION
4. To introduce functions comprise of Census Operation Management for job assignment, eDoc1 Management and supported by multiple modes of data collection (e-Census, CAPI, CATI, PAPI & DOPU) via portal & mobile.

CSS : POST ENUMERATION SURVEY
8. Conduct a sample survey & Platform to estimate the extent of coverage error in the Census, particularly for Living Quarter (LQ) & household (HH)
4. SUPPORT
24 hours with 2 work shifts
Trained Voice Operators
Contact and communication via:
  - Email
  - Hotline
3 tier supports:
  - First Tier: Front end Voice Operators
  - Second Tier: Champions (SME)
  - Third Tier: Technical (IT)
4.2 WAR ROOM

Live updates Dashboards
5. SYSTEM SECURITY
5. SYSTEM SECURITY IN MyIPCS

**APPLICATION**

(Portal, mobile apps, web apps, databases)

- **Code Scan**
  - Source code review analysis as per security best practice

- **Application Scan**
  - Attack simulation to application to discover vulnerability points and security loop holes

- **API Scan**
  - API vulnerability points

- **Mobile App Scan**
  - Identify potential risk which will be exposed through mobile app

**USER GROUPS**

- **External Users**
  - Public Respondents
  - OSL
  - e-CENSUS
  - CT (Job Application)

- **Internal Users**
  - DOSM
  - FL
  - CAPI
  - CATI
  - Portal

**IDENTIFICATION & AUTHENTICATION**

- Username & password protections
- Pictorial security logon ID
- Security questions
- Email verification

**MODULES**

- **Administration**
  - QC
  - CT
  - PES

- **Analytics**

- **Dissemination**

**DATA TRANSFER**

- Secured data transfer methods all incoming external data (Other government agencies, Operation Centers)
- SFTP (Secure File Transfer Protocol)
- Data encryption (when in motion)

**ICT INFRASTRUCTURE**

- Antivirus
- IPS
- SSL
- WAF
6. OUTCOMES
6. OUTCOMES OF ONLINE DATA COLLECTION OF MYIPCS

**CREATE**
- MSAR
- MSPR
- Data Analytic
- 10 integrated online modules

**RAISE**
- Security
- Self-response
- Modernisation in field work

**REDUCE**
- Human resources
- Printing
- Workload

**ELIMINATE**
- Cost
- Workplaces
- Timeline
- Stand alone system

- CREATE
- RAISE
- REDUCE
- ELIMINATE

www.dosm.gov.my
@StatsMalaysia