

The processing of tourists arrivals data



Mr Tshwaro Gae
Ms Duduzile Mazibuko
Tourism Provincial Workshop
25-27 July 2023

IMPROVING LIVES THROUGH DATA ECOSYSTEMS

Outline of the presentation

- Introduction
- Concepts and definitions
- Data sources, scope and coverage
- Process of data acquisition, processing and dissemination
- Limitations and strengths of the statistics
- Conclusions and recommendations

Introduction



History

- Stats SA is responsible for the processing, analysis, write-up and the publication of statistics on international tourists using the data on travellers collected by DHA immigration officers.
- The publications are among the oldest releases and reports in Stats SA dating back to the 1980's.
- Over the four decades, they have gone through several changes regarding
 - definition of key concepts
 - Processing and analysis of data
 - Presentation and dissemination of results

Concepts and definitions



Currently, most of the concepts and definitions used follow the UNWTO recommendations.

- **Traveller** is someone who moves between different geographical locations for any purpose and any duration
- **Visitor** refers to any person travelling to a place other than that of his/her usual environment for less than 12 months and whose main purpose of the trip is other than the exercise of an activity remunerated from within the place visited

- **Same-day visitor** is a visitor who visits a place for less than one night.
- **Overnight visitor** (Tourist) is a visitor who stays at least one night in collective or private accommodation in the place visited.
- **Tourism** comprises the activities of persons travelling to, and staying in places outside their usual environment, for not more than one consecutive year, for leisure, business and other purposes.

Definitions used specifically in-house

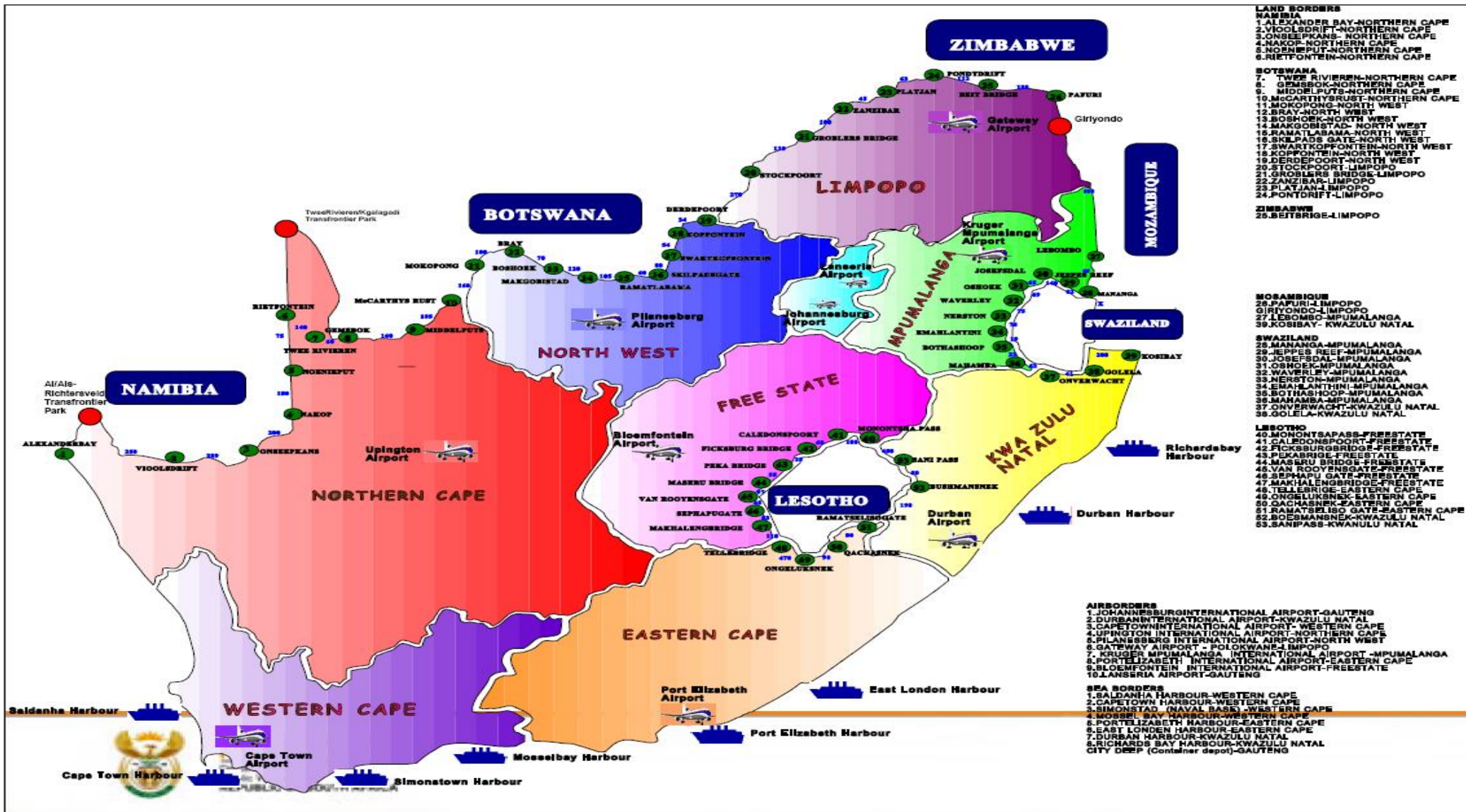
- **South African resident** refers to either a South African citizen or a non-South African citizen with a South African permanent residence permit.
- **Foreign traveller** refers to a traveller who is not a South African citizen or permanent resident.
- **Country** refers to the country of residence of the traveller.

- **Region** refers to Overseas and Africa
 - ✓ **Overseas** sub-regions: Europe, North America, Central and South America, Australasia, the Middle East and Asia.
 - ✓ **Africa** (*refers to all the countries in Africa - both mainland and island countries*) sub-regions: SADC and 'Other' African.

- **SADC** refers to the fourteen countries, *excluding South Africa*, that belong to the Southern African Development Community
- **'Other' African** refers to all non-SADC countries in East and Central Africa; West Africa and North Africa
- **Trip** refers to the arrival and departure of a foreign traveller in a specified month

Data source, scope and coverage





IMPROVING LIVES THROUGH DATA ECOSYSTEMS

- Data on all travellers (South African residents and foreigners) collected by the Department of Home Affairs (DHA) immigration officers at all South African air, land and seaports of entry/exit
- Data collection: electronic – scanning of travel documents and capturing of information directly unto the port's database
- Ports data are transmitted onto the DHA's electronic Movement Control System (eMCS)
- A very crucial characteristic of these border statistics is that though the counts refer to individual persons, statistically the counts are movements of these individuals in a given reference period. Hence the multiple movements of an individual are counted separately

Data file contents

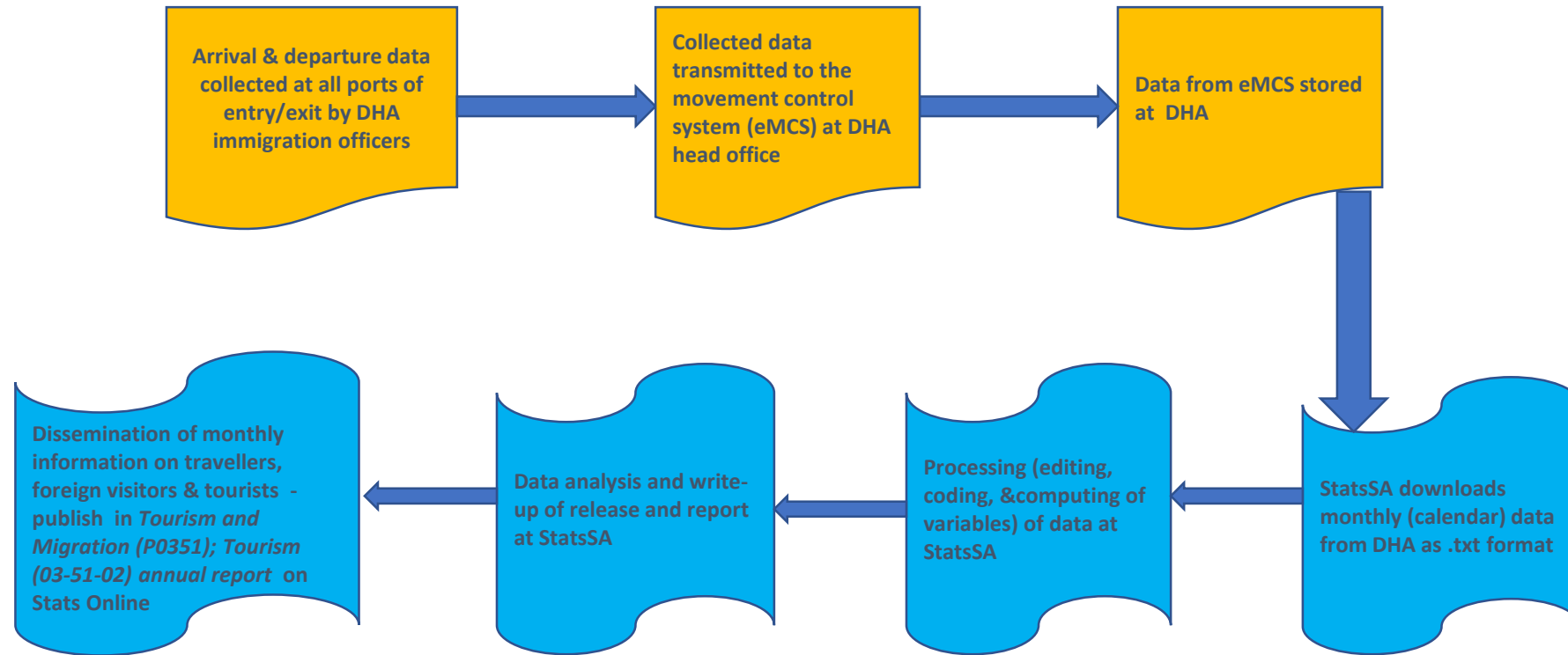
- **Properties of data source files**

- The file is in .txt format
- The first line contains the file's headers (fields)
- Records (values) on every row, including headers, are pipe-delimited

- **Variables**

- All travellers – RSA residents and foreign travellers
- 23 variables from DHA - 16 variables are processed and used for analysis
- Creation of new variables and recoding of old variables

Flow chart on data acquisition, processing, analysis and dissemination on travellers between DHA and Stats SA



Processing



- **Data quality checks and cleaning**
 - Pre-checks: dates, correct year, month and days (1 to 28/29/30/31); data for each day of the month;
 - Check and remove duplicate cases
- **Data confrontation**

The volume of international passengers that were recorded by ACSA at OR Tambo is compared with the immigration data from OR Tambo for each month

Number of ACSA international arrivals compared with DHA international arrivals at OR Tambo international airport, June 2022–June 2023



Analysis



- Classification / categorisation of variables –*e.g.*
 - **country** into sub-regions and sub-regions into regions
 - **age** into age groups;
 - traveller movement reason into purpose of visit
- Descriptive analysis on demographic variables
- Detailed analysis on tourists by country of residence
- Data on South African residents are analysed separately from that on foreign travellers
- All these data can be provided in Excel format to users on request

The bulk of the results are on **foreigners**

- The variables of analysis are: *port of entry, travel direction, mode of travel, purpose of travel, sex, age, visitors (same-day visitors and tourists) and country of residence*
- The country of residence results are further grouped by *sub-region* and *regions*
-

Latest findings



Thank you



statssa.gov.za



twitter.com/StatsSA



facebook.com/StatsSA