

GWG task team on scanner data

Proposed approach, timeline and deliverables for input from team

13 April 2017

Introduction

This is a first attempt (by Frances Krsinich) at defining deliverables, timeframes, and working methods. Any and all feedback from the group is very welcome. This is expected to be a rapidly evolving document.

Frequency and timing of task team meetings

Fortnightly webex meetings from 12 April 2017. We can discuss whether we continue with the current timing (3.30 EST), and/or alternate a different timing that will spread the non-working-hours requirement across the team.

Action: Team to agree on frequency and timing of meetings

Reporting to other groups

We will provide bi-monthly reports to interested groups such as the IWGPS, for their information and to ask for their input and feedback. In particular we need to ensure that we don't duplicate other groups' existing or intended work.

Action: What groups should we be reporting to, other than the IWGPS?

Other members

We plan to ask Alberto Cavallo of the Billion Prices Project to join.

Action: Any other suggestions for new members?

Duration of the project

1 year – ie mid-April 2017 to mid-April 2018.

Iterative approach to delivery

We suggest taking an ['Agile' approach](#). That is, rapid development of a 'minimum viable product', with iterative improvement resulting in increasingly more useful and generalised products.

From this perspective, making the processes already developed by Stats NZ for implementing consumer electronics scanner data in production, sharable and well-documented would be a useful first step. We can then improve and expand from there. The timeframe and deliverables suggested below are based on this approach.

Timeframe and deliverables

Month	Deliverables
mid-April 2017	<ul style="list-style-type: none">• Inaugural meeting of the task team on 12 April
mid-May	<ul style="list-style-type: none">• Contact other international groups such as IWGPS to provide information on the task team's ToR etc• Confirm terms of reference• Agree on (initial) deliverables and timeframe• Invite extra members to join (eg Alberto Cavallo of BPP)• Determine each member's (initial) contributions
mid-June	<ul style="list-style-type: none">• <i>Report to other groups</i>
mid-July	<ul style="list-style-type: none">• R code for FEWS / TDH / GEKS (ie methods for which StatsNZ has SAS)

	<ul style="list-style-type: none"> code already developed) 1st draft of instructional/training guide(s) <ul style="list-style-type: none"> - references to latest relevant literature and the soon-to-be-updated ILO CPI manual's scanner data chapter - instructions on running the R-code developed so far Sharable synthetic data and corresponding example outputs from R-coded methods
mid-August	<ul style="list-style-type: none"> <i>Report to other groups</i>
mid-September	
mid-October	<ul style="list-style-type: none"> <i>Report to other groups</i> R code for basic monitoring and analysis processes (ie based on existing StatsNZ production process for consumer electronics scanner data) R code for extra methods – CLIP (ONS), GK (Statistics Netherlands), CCDI/mean-splice of Diewert&Fox (ABS?), any others? 2nd draft of instructional/training guide(s) Real test data from GfK (consumer electronics scanner data) and PriceStats (webscraped online data) lodged somewhere accessible, such as the Irish sandbox Initial scoping of application platform requirements determined with industry partner(s)
mid-November	
mid-December	<ul style="list-style-type: none"> <i>Report to other groups</i>
mid-January 2018	<ul style="list-style-type: none"> R code fully generalised across <ul style="list-style-type: none"> - multilateral methods (eg FE/TPD, TDH, GEKS, ITGEKS,GK, CCDI) - extension methods (eg movement splice, window splice, half-window splice, mean-splice) - estimation window lengths 3rd draft of instructional/training guide(s) Initial application platform specifications provided to industry partner(s) R code for monitoring and analysis processes finalised
mid-February	<ul style="list-style-type: none"> <i>Report to other groups</i>
mid-March	<ul style="list-style-type: none"> First iteration of application platform tested Updated/clarified specifications for platform delivered to industry partner(s)
mid-April	<ul style="list-style-type: none"> <i>Final report to other groups</i> Application platform available for use by international statistical community Instructional/training guide(s) finalised Example data from GfK (scanner) and PriceStats (online) and/or synthetic data available for use by international statistical community