

RFP 28-2014 - Development of Analytics Capability for PAYE and VAT.

Question	Response
Please can you provide me with a copy of the following documents in response to the Analytics Capability RFP: <ul style="list-style-type: none"> SARS Internal Information Security Systems Life Cycle Policy SARS Information Life Cycle Security Policy 	Document available on the website.
How many SARS fulltime and/or part-time (or extended) resources will be available to work alongside the analytics experts throughout the project (for Knowledge Transfer purposes)?	4-6 Full time resource
Is the PAYE information is on a separate platform from the individual tax return platform?	
What is the number of PAYE references that the analytics need to be run on vs the income tax return numbers?	For an indication please see the Tax Statistics Bulletin with Tender documents has been posted with the Tender documents.
Are you currently performing any analytics and what are these?	yes
At EY, we define analytics as the extensive management and use of data, statistical and quantitative analysis, explanatory and predictive models and fact-based management to improve business performance, drive better business decisions and proactively manage risk. We also define Tax Analytics as the use of analytical techniques and data structuring together with a deep knowledge of PAYE and VAT to identify specific gaps in collection, compliance and optimisation opportunities. Is this also your understanding of Analytics and if not, how do you define analytics for your purposes?	skip
Is your analytics focus rather based on Data Quality or Data Modelling?	Both
Does SARS intend to focus this initiative on the generic analytical component based on collection data OR on analysing the data for fundamental compliance and optimisation scenarios, based on extensive Tax knowledge?	skip
Did we understand you correctly that you only want to have a pilot for PAYE and not for VAT?	yes
Are their business reasons to do PAYE first and do a pilot whereas VAT follows without a pilot? - Would you be able to share why?	skip
Do you currently use external data sources to enrich your master data? Which sources do you use?	yes
If not is there an intention to buy external data to advance your analytics functionalities?	n/a
Data	
What type of information is available within SARS ? Are there any third party sources of data that can be explored (e.g. credit bureau) ?	As an indication, please see the Tax Stats Bulletin that has been posted with the Tender documents. SARS will provide the required 3 rd party data)
How is the information housed in the data systems ? Would it be termed as ready to use without significant cleaning / processing ?	(Access to Data Marts will be provided) Refer to section 8.2.3 and 8.2.4 of the RFP document
Will there be a requirement for setting up database infrastructure to support the required analytics ?	(Infrastructure has been set)
How frequently is the data updated ? Is back in history is the data maintained ?	(Data is mostly updated daily. History data is available)
Capability development :	
What is the existing setup for Analytics capability within SARS ? What kind of skillsets do the existing resources have ?	? Mathematical Sciences, Statistical, Economics, data analysis, Modelling, forecasting, etc.).
Will there be a need to hire and then train the proper resources ?	(Please see the primary objectives of the project, bullet # 4)
Has SARS already decided roughly how many individuals (SARS project team members) it wants to be trained up by this project?	(Please see the primary objectives of the project, bullet # 4 and 8.2.5. of the PRF document), i.e. "Provide an opportunity to work alongside an analytics expert, as well as provide a support system to the SARS team throughout the project, in order to maximise knowledge and skills transfer".
Modeling specific :	
What is the biggest modeling challenge being faced for PAYE / VAT analytics ? Is it related to managing poor / sparse data OR is it related to having un-predictive models ?	skip
Are there any specific population sub-segments which should be targeted as pilot cases for Advanced analytics ?	skip
What is the level of comfort to using machine learning approaches which can lead to extremely predictive models, but can turn out to be extremely complex to understand (black-box)	skip
Is there any requirement to use specific SAS offerings OR can the analysis engine be customized and developed from scratch ?	Use of SAS tools)
Based on our understanding, the key modeling requirement is for files selection models for PAYE and VAT? Is this the case?	Please see 8.3.2)
Would the work include include supporting model for PAYE and VAT? For example assessment of the total earnings of an individual?	skip
Should the analytics work focus on increasing tax revenue, or should substantial weight also be given to other strategic outcomes of SARS' mission (i.e., increased tax fairness by also helping tax payers to properly claim deductions, ease tax filing requirements by reducing documentation requirements, increase cost effectiveness by reducing wasteful compliance/investigation actions)?	Please see Background of the PRF document)

A typical modeling initiative would start with a problem statement and formulation of a modeling approach incl. identification of the required data followed by the extraction of the required data from SARS and other data sources. We have seen extraction of tax filing data and tax investigation outcomes to take anything from 2 weeks to 4 months. Would SARS be able to give any guidance how long a reasonable time budget for data extraction from its data systems would be?	skip
Historical tax investigation data may be of limited use for statistical analysis due to selection bias. Would SARS be open to also create unbiased tax investigation data by auditing randomly selected tax payers during the project duration?	skip
Process	
Did we understand chapter 5 of the RfP correctly that the intended project duration is 12 months?	The successful Bidders will be appointed for a period of 12 months).
Would it be possible to replace the resources allocated to the project in the 12 months period?	refer to 8.2.1)
Would it be possible to use combination of on/off site resources for modeling or do all the resources need to be located at SARS offices?	
The period of 12 months will not cover any support phase. Kindly confirm the understanding	Please see the primary objectives of the project,
'Skill Transfer' refers to Training. Kindly confirm the number of users required to be trained	Please see the primary objectives of the project, bullet # 4 and 8.2.5. of the PRF document), i.e. "Provide an opportunity to work alongside an analytics expert, as well as provide a support system to the SARS team throughout the project, in order to maximise knowledge and skills transfer".
Kindly confirm the tool stack on which the trainings would required to be given	See 1 above
Kindly confirm the number of days of training required to be given	See 1 above
Kindly confirm whether any travel would need to be undertaken to conduct the trainings	See 1 above and 8.2.1
Kindly confirm whether all infrastructure required for the training would be provided by SARS or would the bidders be required to provide the same	The infrastructure is in place, bidders will bring in their laptops
Kindly mention the number of source systems that would need to be integrated	skip
Kindly provide the following details on the source systems that would need to be integrated :	
- Name	
- Data Volume	
- % Growth in volume (y-o-y)	See tax statistics bulletin
- Application Platform	
- Backend Database	
- Server Hardware Application	
How much historical data is expected (no. of years?, data size?)	Minimum 5 years
Would the entire team (including the development team) be required to be located out of SARS premises in Pretoria or can they be allowed to be located out of the bidders premises in an onsite-offshore model	8.2.1
Can Phase 1 and Phase 2 be carried out in parallel or does it need to be done on a sequential manner only	Please see Background "The development of VAT Analytics Capability will follow a similar approach to the PAYE Analytics Capability and incorporates use of the knowledge and skills gained from the PAYE Analytics Capability project."
Kindly provide an estimate of the number of 'measures' and number of 'dimensions' for each of the phases respectively	skip
Kindly provide an estimate of the number of reports required to be built for each of the phases respectively	skip
Kindly provide an estimate of the number of dashboards required to be built for each of the phases respectively	skip
Are there existing data models used for PAYE and VAT analytics? Should these models be considered as part of business understanding phase?	
What is considered as 'large data sets'	skip
How many projects need to be mentioned for this clause	skip
Will international citations be considered as fulfilling criteria? ?	skip
How many projects need to be mentioned for this clause	skip
Is this required for resources with SAS skill set only or is it required for all the layers of the solution (Database, ETL, Reporting, Analytics)	skip
Kindly confirm whether these can be global references	Yes
The weightages against the Sections seem to be subjective. Kindly provide details of the specific parameters with weights against each of the line items in the schedule for a better understanding of the scoring system	skip
It is assumed that all the infrastructure required for the solution including the environments would be provided by SARS. Kindly confirm	Yes. Bidder to provide laptops
It is assumed that the necessary infrastructure would be in place from day 1 of the assignment. Kindly confirm	Yes
Kindly mention the environments to be considered for the solution (Development, Testing, Production) if this needs to be provided by the bidders	skip
Kindly mention the minimum hardware sizing requirement to be considered for the environments if this needs to be provided by the bidders	skip

Kindly mention the minimum hardware requirements for all individual infrastructure items required to be supplied by the bidder	skip
Has the entire solution stack (Database, ETL, DQ, Reporting, Analytics) already been decided and implemented.	skip
If the answer to the above question is yes, kindly provide the details of the solution stack so chosen / implemented.	skip
Would bidders be required to procure and provide the solution stack to SARS	skip
If the answer to the above question is yes, kindly provide the number of users : - Normal Users - Power Users - Concurrent Users for the proposed solution. Kindly provide separate numbers for PAYE and VAT respectively	
Kindly mention the growth in the number of users expected over the next 3 years	skip
Kindly mention the ATS period to be budgeted in the estimates if the solution stack is to be provided by the bidders	skip
Billing Milestones not mentioned in the RFP. Kindly share the billing milestones proposed for the project	Indicated in pricing template
What is the underlying existing architecture for PAYE and VAT applications , if existing already	skip
What is the SAS solution components to be customised for this project	skip
Can we have both PAYE and VAT analytics implementation tracks in parallel? Or there are interdependency from source data sets perspective. Pl clarify.	No
Does SARS have an existing SAS Analytics: (1) Framework (2) Team and (3) Environment. Please give breakdown / details of each of the above.	yes