ADOA-ASET

Arizona Strategic Enterprise Technology

Project Investment Justification

Version 01.01

A Statewide Standard Document for Information Technology Projects

Project Title:

SAS Software for Improved Statistical Data Analysis

Agency Name:	**Arizona Department of Administration's Office of Employment & Population Statistics**
Date:	**December 2, 2013**
Agency Contact Name:	**Vignesh Sukumaran**
Agency Contact Phone:	
Agency Contact Email:	

[Hover for Instructions](#)

I. Management Summary*

The Arizona Department of Administration's Office of Employment & Population Statistics (ADOA-EPS) employs sophisticated researchers whose responsibility is to explore all aspects of the Arizona labor market as well as the demographic conditions of the state. They are proficient programmers and experienced users of computing software who work with extremely large datasets. The main statistical computing software currently in use is IBM SPSS. Its weak performance in data management and programmability has caused researchers to migrate to alternate solutions that are inefficient.

The proposed solution, SAS software, offers an acclaimed fourth-generation programming language and powerful tools for data management. Based on conversations with in-house SAS experts and a SAS systems engineer, it is proposed that ADOA-EPS researchers be provided specific SAS software modules. This investment in better analytics software will result in improved process development, code reuse, collaboration and a demonstrable increase in speed with which data analysis is done.

II. Project Investment Justification (PIJ) Type*

☐ Yes ☒ No   Is this document being provided for a Pre-PIJ / Assessment phase?

If Yes,

Identify any cost to be incurred during the Assessment phase.	$
Based on research done to date, provide a high-level estimate or range of development costs anticipated for the full PIJ.	$

Explain:
Click here to enter text.

☐ Yes ☒ No   Will a Request for Proposal (RFP) be issued as part of the Pre-PIJ or PIJ?

III. Business Case

A. *Business Problem**

The number of stakeholders and the complexity of analysis are being constrained because ADOA-EPS does not have the advanced analytical tools that are available. To ensure the highest quality analysis, processing and computing needs have to keep pace. Bottlenecks such as inefficient software solutions will have to be removed. IBM SPSS (SPSS) is currently used in ADOA-EPS as the main statistical computing software and has the following limitations and problems:

1) Data management, such as querying and retrieving data from relational databases, and manipulation of large datasets has been difficult. In certain projects, SPSS could not handle large data files and workarounds had to be devised to provide a solution resulting in project delays.

2) SPSS syntax and the programming interface have remained counterintuitive and difficult to use even after the release of many new versions and acquisition of the company by IBM in 2009.

B.   *Proposed Business Solution**

Analyst positions in ADOA-EPS are filled with researchers experienced in the use of SAS software. Current business problems need cutting-edge statistical procedures, good data management and a robust programming environment for process automation. So, to ensure business continuity and success in data analysis, this proposal recommends SAS software for the desktop.

C.   *Quantified Benefits**

X	Service enhancement
	Increased revenue
	Cost reduction
X	Problem avoidance
	Risk avoidance

Explain:

SAS will provide much needed features that are either unavailable or difficult to use in SPSS. Some of the immediate benefits are outlined here. This list is not exhaustive and is compiled from discussions with researchers in ADOA-EPS who are experienced SAS users.

1) SAS's programming environment is powerful and well-developed. The equivalent SPSS programming interface creates syntax that is overly complicated and unintuitive.

2) SAS is renowned for its data management. Its parallel-processing ability allows simultaneous processing of numerous data files. Unlike with SPSS, data files don't need to be opened before use, which improves use of I/O and memory.

3) SAS has a highly optimized data access engine for work with relational databases. This includes high performance features such as streaming results, buffering, compressions, threading, etc. Support for temporary tables, control of join-processing, and access to stored procedures from programs are some of the other features that are available in SAS and not in SPSS.

IV.   Technology Approach

A.   *Proposed Technology Solution**

Install Base SAS, SAS/STAT, SAS/GRAPH, SAS/ETS and SAS/ACCESS software on the desktop of 12 research analysts.

B.   *Technology Environment*

Research analysts in ADOA-EPS use Hewlett Packard Z Series workstations with Intel Xeon processors and RAM memory ranging from 8 GB to 24 GB.

The research analysts (3 Ph.D. and 7 master's degree holders) have used SAS in graduate studies and in prior workplaces. They are proficient programmers who use Python[1], VBA[2] and R.[3]

Data is mostly stored and processed using relational databases. These databases are managed by a database administrator in ADOA-EPS on servers hosted by the ADOA Data Center. Data files with structured data[4] are also in use.

C.   *Selection Process*

Among statistical computing software, SAS, SPSS and Stata are prominent. Skill sets, cost, deployment models and feature requirements were reviewed.

1) SAS provides a powerful computing platform that outshines the other packages in data management and programmability. Data management and programmability are two very important considerations ADOA-EPS used to distinguish available software products.

2) University graduates in disciplines such as economics, mathematics, statistics, etc., are used to fill positions in the office. Real-world classroom projects rely on SAS, R and other software. Experience and knowledge in SAS is regarded as a job-market advantage. The probability of hiring researchers with SAS experience is high.

3) The SAS client-server model is expensive. A sample quote for an entry level client-server solution based on a server with a single quad-core processor is over $125,000. This does not include server hardware, ongoing server support, server administration and one-time server administration training costs.

A peripheral consideration, which is important nevertheless, is the use of SAS by our federal partners; Bureau of Labor Statistics (BLS), Employment and Training Administration (ETA) and Census Bureau. Using common software allows better exchange of ideas, data and programs between federal-state partners. For example, a widely used seasonal adjustment[5] program called "X-13ARIMA SEATS," created by the Census Bureau, has accompanying graphical diagnostic software that is based on SAS/GRAPH. If provided SAS modules identified in section IV-A, analysts in ADOA-EPS can use the diagnostic software created by the bureau without the need for reinvention.

V.   Project Approach

A.   *Project Schedule**

Project Start Date: 12/16/2013   **Project End Date**: 2/21/2014

B.   *Project Milestones*

Major Milestones	Start Date	Finish Date
Purchase of SAS software using a PO	Dec-16, 2013	Jan-31, 2014
Installation of SAS software on 12 desktops	Jan-31, 2014	Feb-21, 2014

VI.   Roles and Responsibilities

A.   *Project Roles and Responsibilities*

Project Manager: Vignesh Sukumaran
Technical Lead: Gilbert Adongo (SAS certified analyst in EPS)
SAS Systems Engineer: Timo Kettunen
SAS Account Executive: Sarah Wharton
SHI Account Manager: Kunal Patel
The project manager will work closely with the ADOA-EPS technical lead, SAS engineer and ADOA-ASET End User Support Team during the product acquisition and installation phases. Betty Camarena in ADOA-EPS will work with Joyce Wright to complete the PO process.

B.   *Project Manager Certification*

	Project Management Professional (PMP) Certified
	State of Arizona Certified
X	Project Management Certification not required

C.   *Full-Time Employee (FTE) Project Hours*

Total Full-Time Employee Hours	40
Total Full-Time Employee Cost	$

VII.   Risk Matrix, Areas of Impact, Itemized List, PIJ Financials

VIII. Project Approvals

A. Agency CIO Review*

Key Management Information	Yes	No
1. Is this project for a mission-critical application system?		X
2. Is this project referenced in your agency's Strategic IT Plan?		X
3. Is this project in compliance with all agency and State standards and policies for network, security, platform, software/application, and/or data/information as defined in http://aset.azdoa.gov/security/policies-standards-and-procedures, and applicable to this project? If **NO**, explain in detail in the "XI. Additional Information" section below.	X	
4. Will this project transmit, store, or process sensitive, confidential or Personally Identifiable Information (PII) data? If **YES,** in the "XI. Additional Information" section below, describe what security controls are being put in place to protect the data.		X
5. Is this project in compliance with the Arizona Revised Statutes (A.R.S.) and GRRC rules?	X	
6. Is this project in compliance with the statewide policy regarding the accessibility to equipment and information technology for citizens with disabilities?	X	

B. Project Values*

The following table should be populated with summary information from other sections of the PIJ.

Description	Section	Number or Cost
Assessment Cost **(if applicable for Pre-PIJ)**	II. PIJ Type - Pre-PIJ Assessment Cost	$
Total Development Cost	VII. PIJ Financials tab	$62,488.01
Total Project Cost	VII. PIJ Financials tab	$126,081.77
FTE Hours	VI. Roles and Responsibilities	40

C. Agency Approvals*

Contact	Printed Name	Signature	Email and Phone
Project Manager:	Vignesh Sukumaran		
Project Sponsor:	Paul J. Shannon		
Agency Information Security Officer:	Mike J. Lettman		
Agency CIO:	Aaron V. Sandeen		
Agency Director:	Brian C. McNeil		

IX.   Optional Attachments

SHI Quotation # 7279634

SHI Quotation # 7295238

X.   Glossary

1) Python – Open source programming language (www.python.org)

2) VBA – Visual Basic for Applications is a programming language used to extend the standard capabilities of Microsoft's Office software.

3) R – A programming language and an interactive environment for doing statistics.

4) Structured data – data from web pages, email, social media etc. are unstructured and are in contrast to structured data that is in columns and rows.

5) Seasonal adjustment - Seasonal adjustment is a statistical technique that attempts to measure and remove the influences of predictable seasonal patterns to reveal how employment and unemployment change from month to month in the programs managed by ADOA-EPS.

XI.   Additional Information

X-13ARIMA-SEATS Seasonal Adjustment Program page:
http://www.census.gov/srd/www/x13as/. This is referenced in section IV-C.

Links:

ADOA-ASET Website

ADOA-ASET Project Investment Justification Information Templates and Contacts

Email Addresses:

Strategic Oversight

ADOA-ASET_Webmaster@azdoa.gov