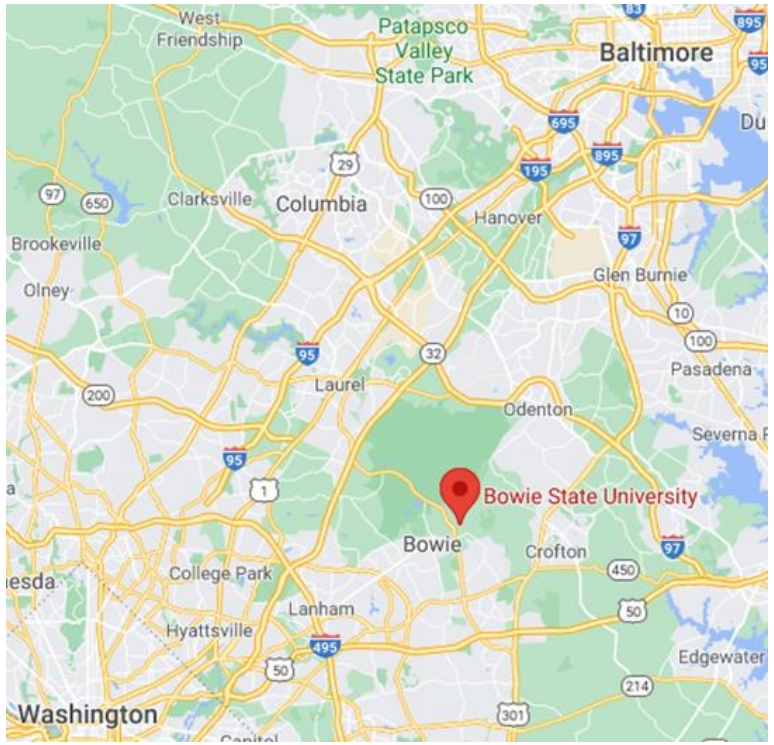
# University Enrollment Analytics with OAC on ADW

Tim Vlamis, Vlamis Software Solutions

# Bowie State University




- 150+ years old
- Oldest Historically Black College and University in Maryland
- Located between Washington DC and Baltimore MD
- Part of the University System of Maryland
- 23 undergraduate, 19 master's, 2 doctoral programs
- Enrollment – over 6,300
  - ~5,400 UNG
  - ~900 GRAD
- Over 1,000 degrees awarded annually



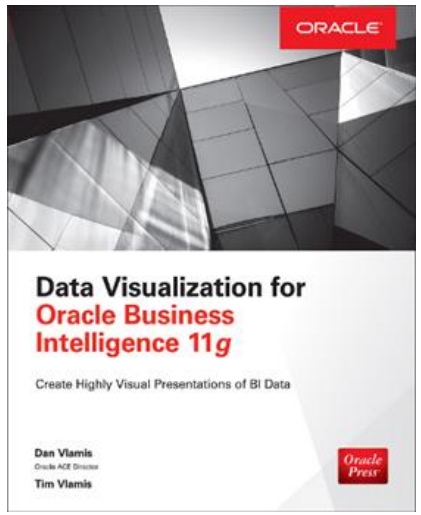
vlamis™
SOFTWARE SOLUTIONS

# Vlamis Software Solutions

- Vlamis Software founded in 1992 in Kansas City, Missouri
- Developed 400+ Oracle BI and analytics systems
- Specializes in Oracle-based:
  - Enterprise Business Intelligence & Analytics
  - Analytic Warehousing
  - Machine Learning and Predictive Analytics
  - Data Visualization
  - ETL and data integration
- Multiple Oracle ACEs, consultants average 15+ years
- Creators of the Force Directed Graph Plugin on Oracle Analytics Library
- www.vlamis.com (blog, papers, newsletters, services)
- Co-authors of book "Data Visualization for OBI 11g"

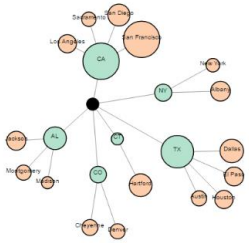
ORACLE | Service Partner

*Expertise in*
**Oracle Cloud Platform**
**Business Analytics**
in North America




ORACLE ACE Director

ORACLE ACE Director





vlamis™
SOFTWARE SOLUTIONS

# Office of Planning, Analysis and Accountability (OPAA)

IR Responsibilities:

- Data Reporting (federal, state, system)
- Enrollment analysis
  - Student retention, progression, graduation, student profiles, fact book
- Grade analysis
  - Midterm analysis, success rates, etc
- Surveys
  - **Internal** - entering student, current student, graduating student satisfaction, course evaluations, ad hoc
  - **External** –US News, CUPA, AAUP, VSE, etc
- Enrollment Projections

IE Responsibilities:

- Narrative Reporting (state, system)
- NCAA / USDOE reporting
- Curriculum management
- Accreditation – regional / specialized
- Strategic planning
- Planning & budgeting
- SLO and other assessment activities
- President and Provost projects
- ***Data warehouse and Post Graduation outcomes***



vlamis™
SOFTWARE SOLUTIONS

# Analytics Development in 2 Phases

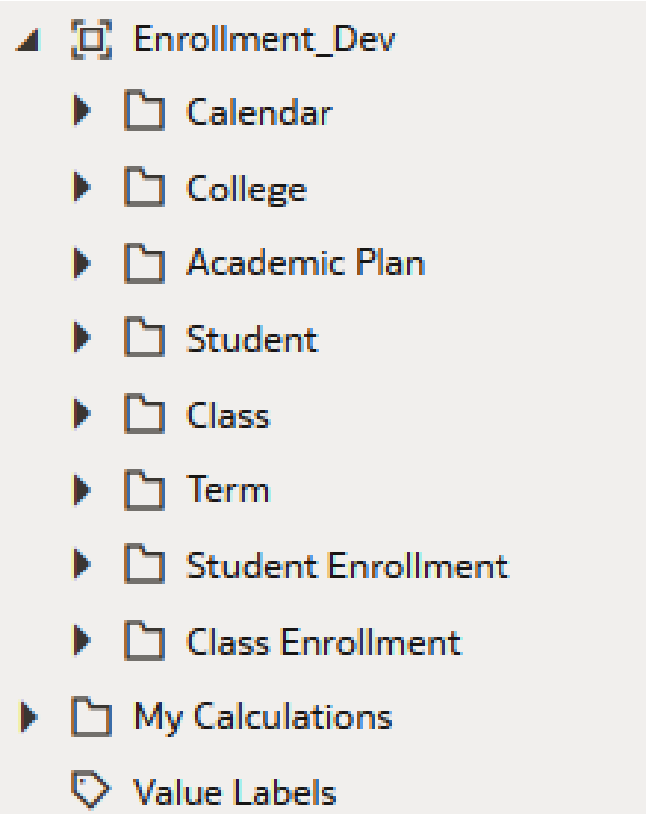
- BSU was using a combination of direct SQL queries, PeopleSoft reports (Campus Solutions), Tableau, and lots of MS Excel
- Oracle proposed OAC and ADW combo
- BSU and Oracle brought in Vlamis
- Phase 1 – Demonstrate quick results with Postgrad Analytics
- Phase 2 – Design and implement Enrollment Analytics on ADW

# Different Approaches for Projects

- Post grad analytics used ADW and OAC data sets with DV
- Enrollment analytics used ODI from PeopleSoft to ADW
  - Developed Enrollment Subject Area in OAC repository
  - Compared Classic vs. DV front end
- Trade off of speed, expense, complexity, robustness and longevity

# Enrollment Analytics

- Uses ODI for ETL from PeopleSoft on-prem to ADW
- ADW to OAC repository
- Enrollment Subject Area designed from scratch
- Client chose DV interface over Classic






vlamis™
SOFTWARE SOLUTIONS

# Enrollment Analytics

- Collaborated with BSU and Sierra Cedar SMEs
- Determined initial set of PeopleSoft tables
  - Most important was use of PS_STDNT_ENRL and not PS_STDNT_CAR_TERM for enrollment fact table
  - Enrollment events, not current state of student enrollment by career
- Dimensional, hierarchical view of data different than transactional state view of data
- Does not include admission, grades/progress, degree, postgrad
- Individual student headcount enrollments AND class enrollments
- Academic years vs calendar years



vlamis™ SOFTWARE SOLUTIONS

# Enrollment Data Challenges

- Lots of PeopleSoft data inconsistencies
- Academic year vs calendar year
  - (settled on eliminating academic year 2022 to "catch up"
- Class sessions and naming inconsistencies
- Latest change date used to determine slowly changing dimension values (BSU reporting not always consistent)
- Cohort and status determination challenges (first enrolled date)

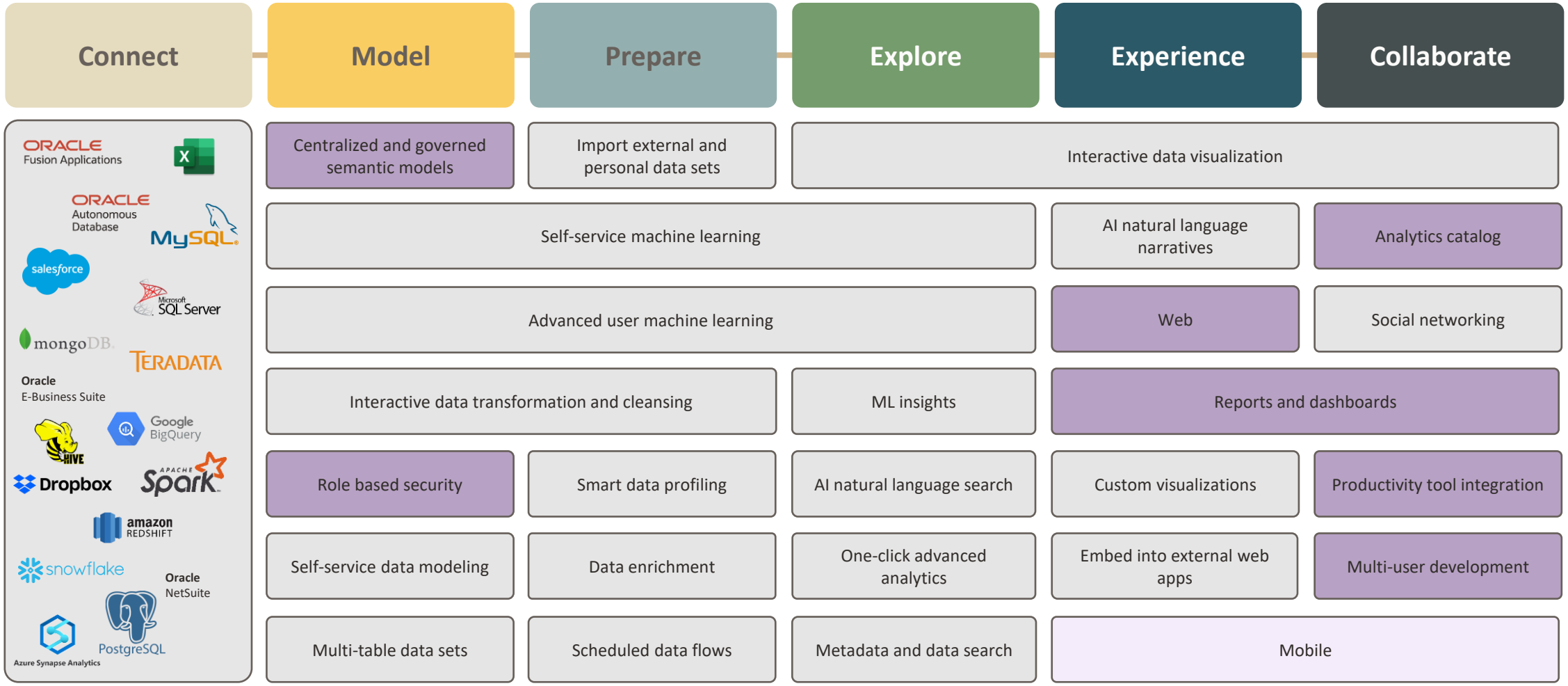
# Classic vs Data Visualization Demo

- Develop examples on both and demo pros and cons
- BSU stakeholders see progress on project
- Discuss classic vs DV front end.
- Ultimately make decision on deployment before invest too much time
- Receive feedback on options in deployment
  - Prompt placement and style
  - Naming and titles
  - Graph and table design
  - Color palettes and assignments – any BSU standards?
  - Key messages in first dashboard page?  Doesn't update much daily?

# Good Questions Guide Priorities

- What is the key message or insight?
- Which is more important, this or that?
- What comparison do you want? Difference? Percentage difference? Etc.
- This version emphasizes this, this other version emphasizes that, which do you prefer?
- What do you want everyone else to understand from this?
- Would you rather we spend time on this or that?
- Who else in your team will use this? Will others outside your team use this version?
- Currently we have x hours estimated to work on this, which is (high, low, average), sound right?

# Oracle Analytics addresses the needs of the entire analytics workflow






vlamis™ SOFTWARE SOLUTIONS

# DV things to demo

- Create custom calculation
- Layout of screen is drag and drop – WYSIWYG, free style layout
- Use grammar pane to rotate a dimension, add a column
- Change type of visualization
- Formatting tables, pivot tables
- Trellising – in all vizes
- Property pane – in all vizes
- Filtering – dashboard level (pinned), canvas level (up top), viz
- Brushing, selection, right-click

# DV Bonuses

- Data preparation
- Data flows
- Machine learning

# Classic things to demo

- Create custom calculation
- Edit dashboard – UI is classic table presentation of screen
- Dashboard is made up of analyses-queries and views in layouts
- Add different visualization to layouts
- Sections can be expanded/collapsed and based on conditions
- Formatting tables, pivot tables
- Prompts are a separate object – reusable
- Right-click menus – data actions
- Selection steps – extra data filtering

# Answers and Dashboards SWOT

## Strengths

- Highly designed dashboards
- Query definition power
- Prompts and selection steps
- Good control of table formats
- Decent selection of graphs
- Hierarchical columns

## Opportunities

- Ability to set standards

## Weaknesses

- Dynamic layout WYSIWYG
- High training costs
- Endlessly deep menus
- Poor dashboard layout tool

## Threats

- Visual analytics and data discovery tools
- Poorly designed repositories



vlamis™
SOFTWARE SOLUTIONS

# Data Visualization SWOT

## Strengths

- Extremely interactive
- Framework for brushing
- Mashup
- No RPD required
- Data source connections
- Leverages Oracle security

## Opportunities

- Machine learning is promising
- Data Flows transform data

## Weaknesses

- Highly designed dashboards
- Weak documentation

## Threats

- Starting from behind
- Users already love Tableau and PowerBI

# Advantages of DV Interface

- Area of Oracle investment
- More modern interface
- Simpler interface for ad-hoc use
- More interactive – e.g. brushing
- Competes with Tableau and other more modern BI tools
- Ability to use external data – NO RPD mapping required!
- Mash up with external data (including spreadsheets and cloud)
- Integrated machine learning, AI, and NLP (natural language)
- Extend with custom visualizations
- Emphasis on self-service

# Extra DV Features

- New conditional formatting primitives including ordering
- Mobile-ready
- Embeddable in other web applications
- Data preparation and data flows wrangle data
- Machine learning primitives
- Auto Insights.  Can tell user where the story is.
- Adhoc exploration of data not defined by RPD

# Reasons for Classic Answers / Dashboards

- ~~Investment in current dashboards and reports~~
- Answers allows more control – ~~e.g. font control~~
- Highly customized reports – custom queries
- Prompts have different capabilities (exist as prompt objects)
- More control over layout, especially of tables
- Can use repository variables (not in DV yet, but soon)
- Selection steps and custom groups – post-aggregation

# Post Graduate Outcomes Project

- BSU Strategic Plan – Objective 3.4
  - Construct an analytics capacity (first to launch)
- Title III HBGI Activity
  - Enhancing Institutional Effectiveness, Academic, and Student Success Through Data Analytics
    - ✓Phase I – Identifying data warehouse and analytics tools
      - Oracle Autonomous Data Warehouse (ADW) & Oracle Analytics Cloud (OAC)
    - Phase II – Build initial data structures to support academic decision making and student success
      - SOW 1 – Create initial data integrations from Peoplesoft to ADW
      - ✓SOW 2 – Create data structures for tracking reenrollment of bachelor's degree recipients



vlamis™
SOFTWARE SOLUTIONS

# Tracking Re-enrollment of Bachelor's Degree Recipients

- Why
  - Demonstrate effectiveness of BSU's academic programs
    - Post Graduation Reenrollment
    - Post Graduation Earnings (future)
  - Meet programmatic specialized accreditation requirements
  - Create a sustained and accessible data integration and visualization process

# Oracle Technologies Involved

- Oracle Autonomous Database (ADW)
  - Loaded data into ADW via ADW Database Actions (“ADW Tools”)
  - Transformed data using SQL Developer – SQL UNPIVOT
- Oracle Analytics Cloud (OAC)
  - Used Data Flows to create hierarchy of codes
  - Used Multi-table dataset to access data directly from ADW
  - Using RPD for enrollment analytics, but not for this postgrad project

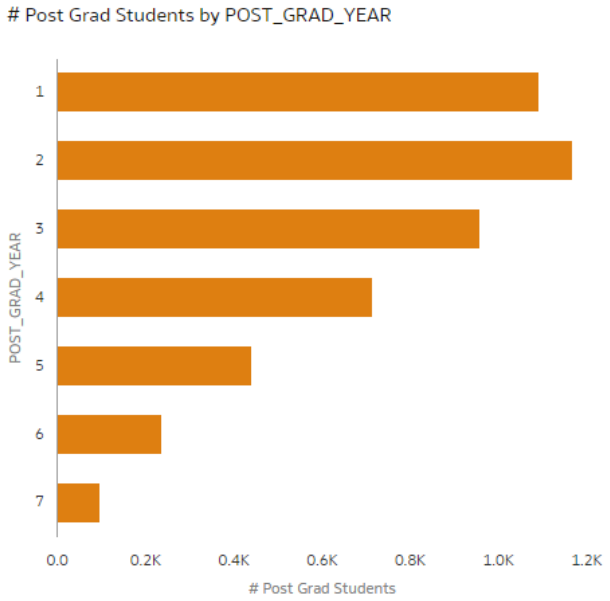
DEGREE_SOUGHT All
POST_GRAD_YEAR <NULL>
COLLECTION_YEAR 2014

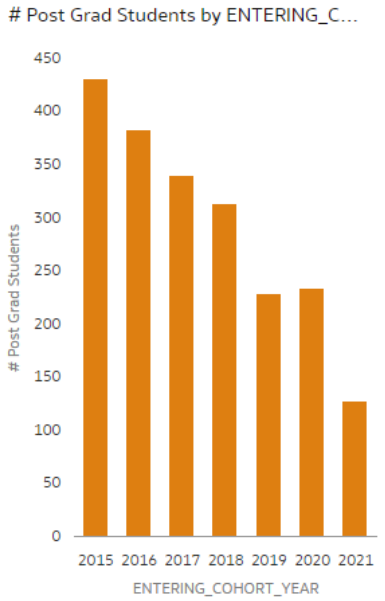
DEGREE_SOUGHT: 40
PROGRAM_NAME: All
Pattern of First Enrollments ba...: All
ENTERING_COHORT_YEAR: All
Data Year: All
Name of First Different Four-Ye...: All
ETH_97: All

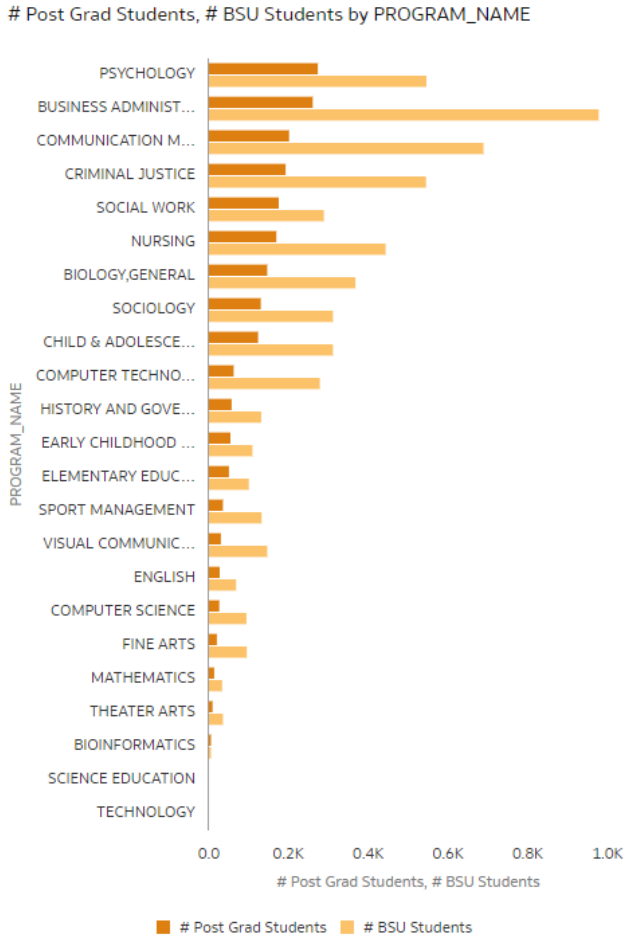
# Post Grad Students Since 2015
2,056

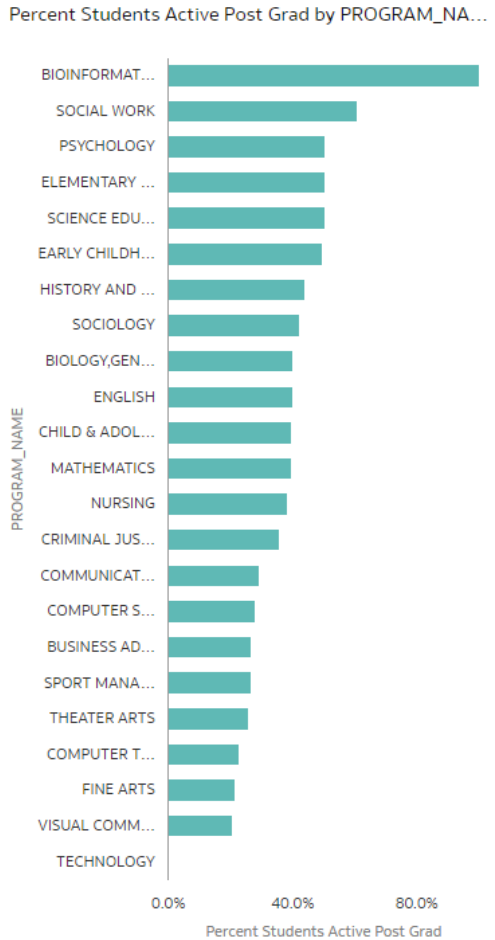
Percent Students Active Post Grad
36.1%

# BSU Students
5,698



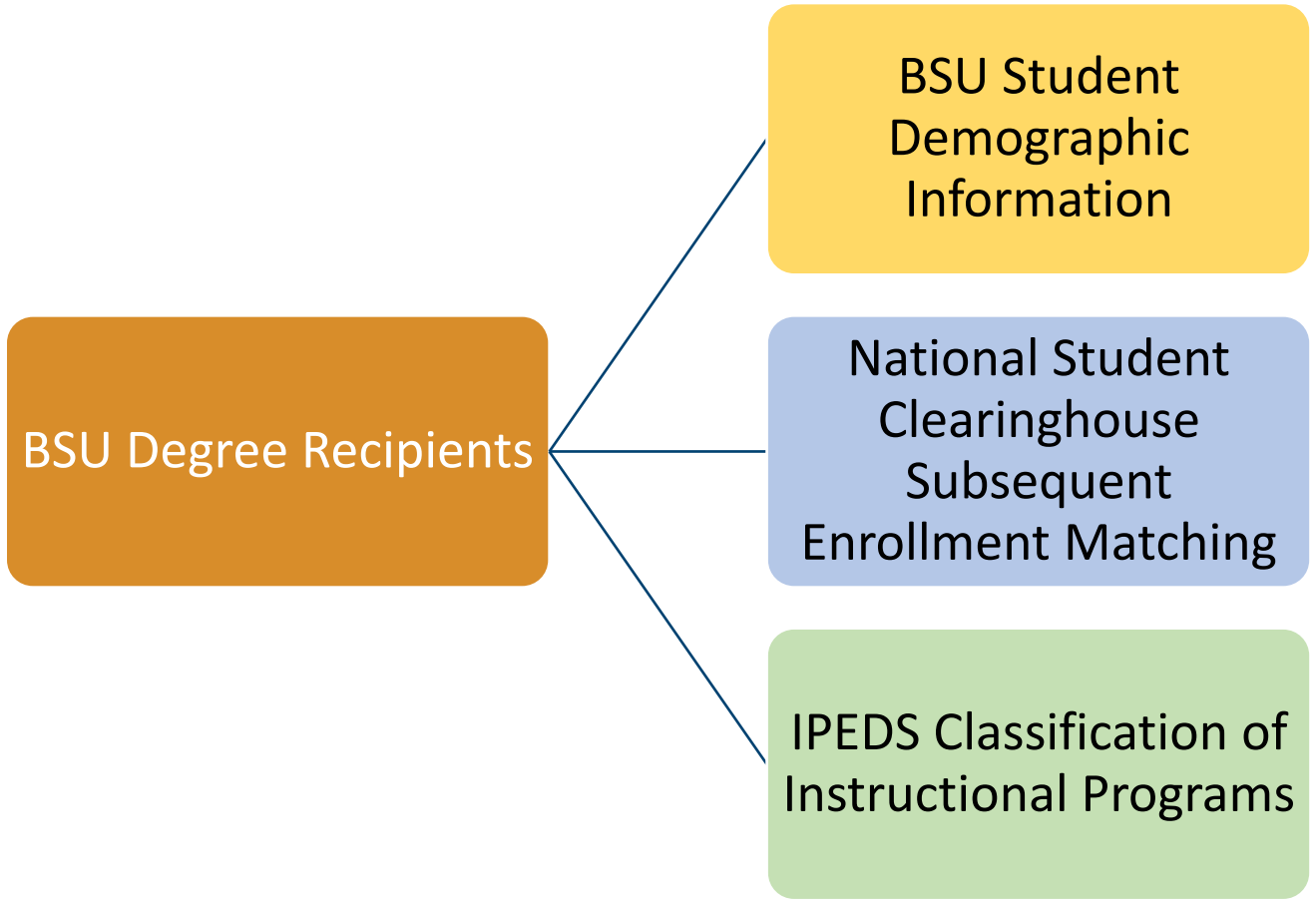

ENTERING_COHORT_YEAR, # NSC Students, POST_GRAD_YEAR


# NSC Students by ENTERING_COHORT_YEAR, POST_GRAD_YEAR

First 4Y Inst With BSU: <NULL>

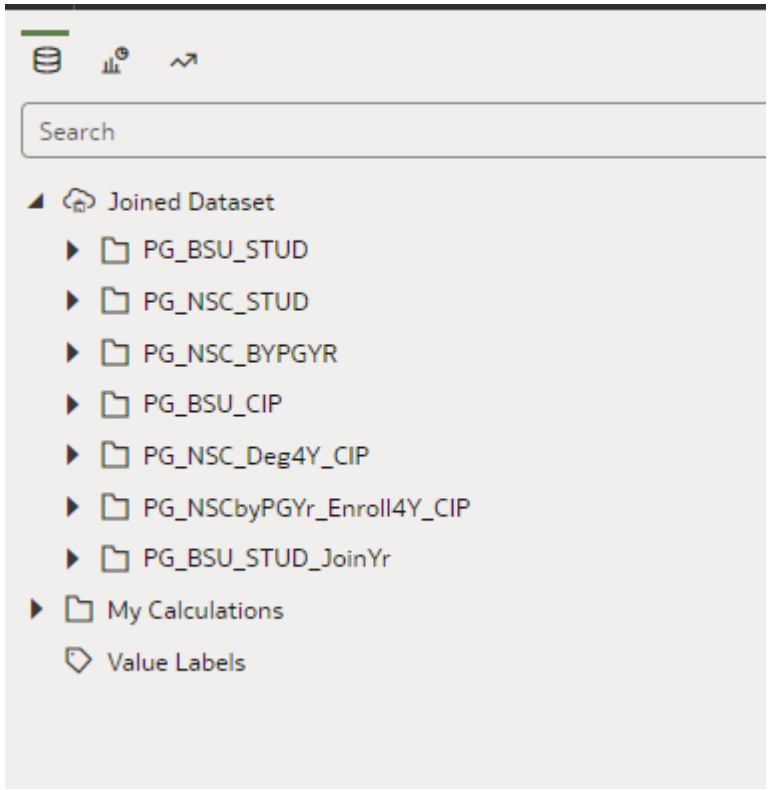
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Grand Total |
|---|---|---|---|---|---|---|---|---|
| | # NSC | # NSC | # NSC | # NSC | # NSC | # NSC | # NSC | |

# Tracking Reenrollment of Bachelor's Degree Recipients – **Data Sources**



# Data Strategy for Post Grad

- Trade off - **development speed** vs quality
  - Used multi-table data sets rather than RPD
  - RPD development and ADW base model not complete
- Had to consider modeling in ADW vs OAC
- Custom calculations to be done in front end
- Move modeling “down the stack” in the  future





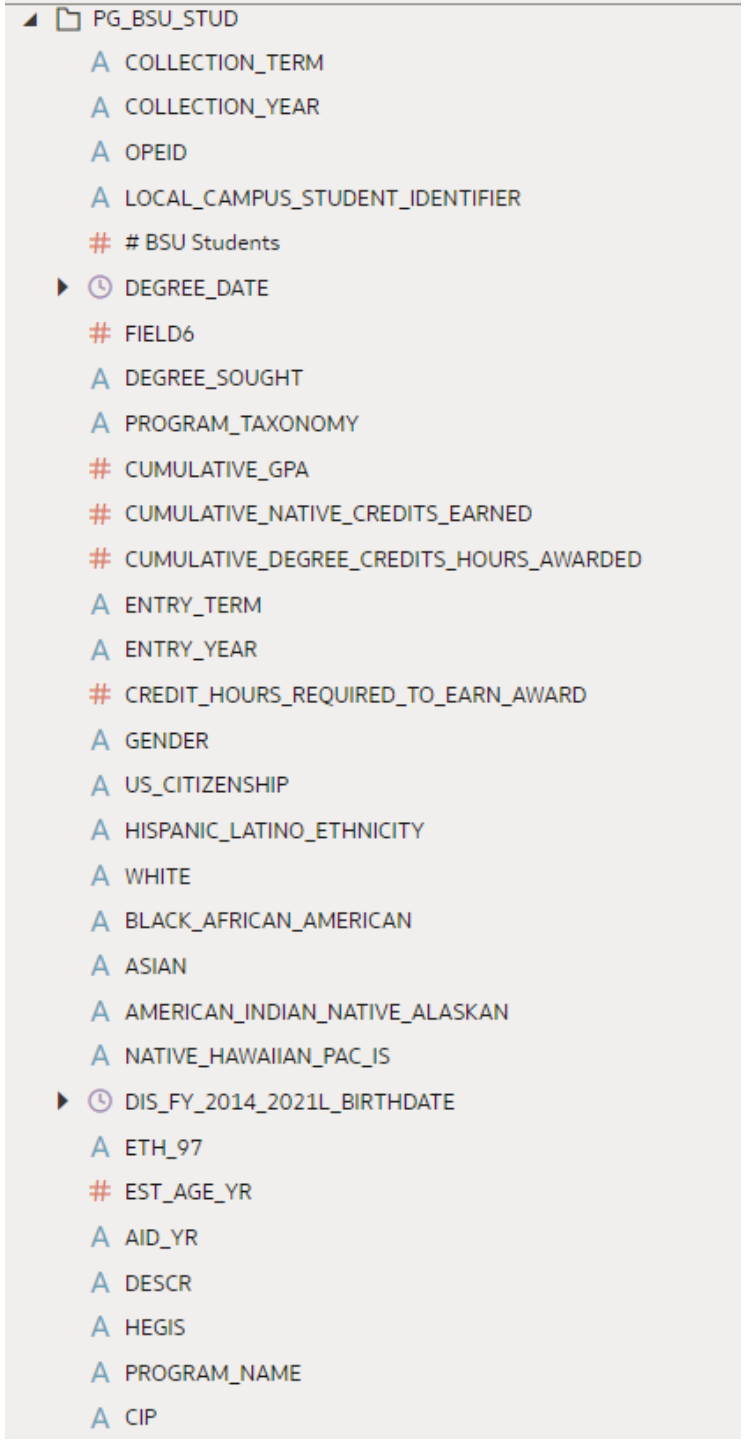

vlamis™
SOFTWARE SOLUTIONS

# Data Sources – Ad Hoc and Warehouse

- Created summary file for BSU students
- Loaded data into ADW using Database Tools utility
  - Decided to use power of ADW vs OAC data sets
  - 3 major sources – BSU, Clearing House, CIP
- Used staging tables to facilitate unpivot transforms for clearing house data
  - Post grad years 1 – 8 table PG_NSC_byPGYr
  - Creates 8 records per student, one for each year
- Had to clean up and create short titles for CIP data

# Data Sources – Student Data

- Main Student Table from BSU has one record per student
- Compiled data from PeopleSoft tables

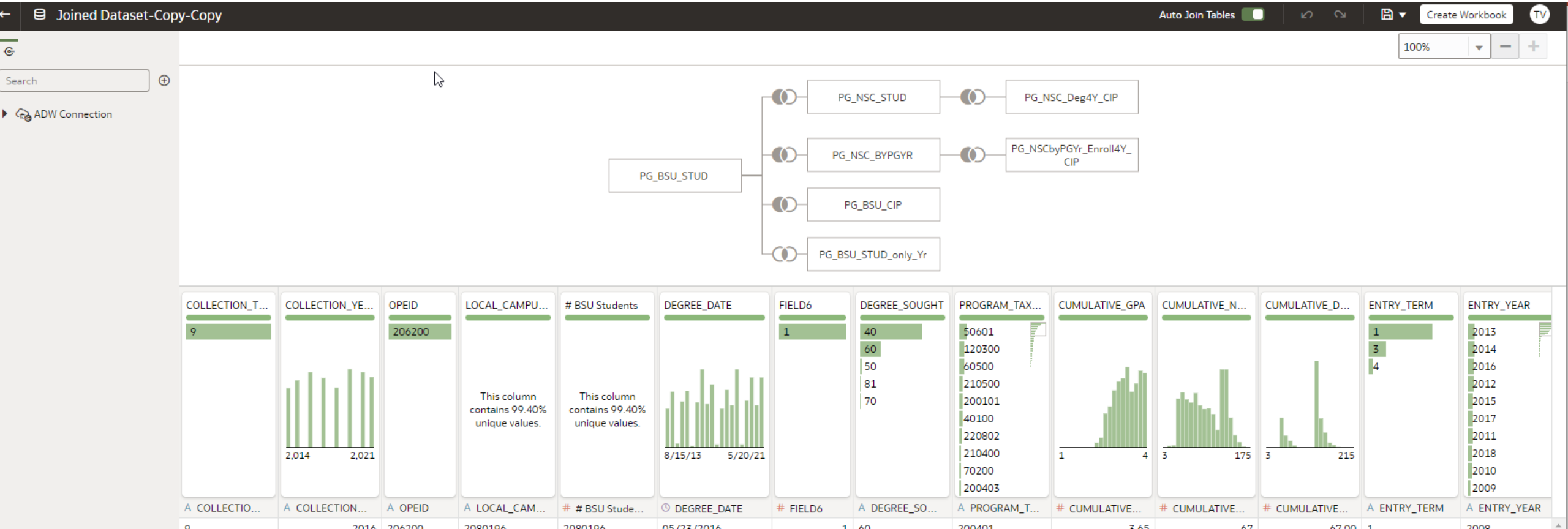





vlamis™
SOFTWARE SOLUTIONS

# Data Sources – CIP Codes –Topics of Study

- CIP Codes maintained by IES National Center for Educational Statistics
- Updated in 2010, 2020 “Crosswalk”
- Website listing https://nces.ed.gov/ipeds/cipcode/browse.aspx?y=55
- “There’s a code for that.”
- Classification of Instructional Programing


Joined Dataset-Copy-Copy
Auto Join Tables
Create Workbook
TV
100%
Search
ADW Connection
PG_NSC_STUD
PG_NSC_Deg4Y_CIP
PG_BSU_STUD
PG_NSC_BYPGYR
PG_NSCbyPGYr_Enroll4Y_CIP
PG_BSU_CIP
PG_BSU_STUD_only_Yr
COLLECTION_T...
COLLECTION_YE...
OPEID
LOCAL_CAMPU...
# BSU Students
DEGREE_DATE
FIELD6
DEGREE_SOUGHT
PROGRAM_TAX...
CUMULATIVE_GPA
CUMULATIVE_N...
CUMULATIVE_D...
ENTRY_TERM
ENTRY_YEAR
9
2,014
2,021
206200
This column contains 99.40% unique values.
This column contains 99.40% unique values.
8/15/13
5/20/21
1
40
60
50
81
70
50601
120300
60500
210500
200101
40100
220802
210400
70200
200403
1
4
3
175
215
1
3
4
2013
2014
2016
2012
2015
2017
2011
2018
2010
2009
COLLECTIO...
COLLECTION...
OPEID
LOCAL_CAM...
# BSU Stude...
DEGREE_DATE
FIELD6
DEGREE_SO...
PROGRAM_T...
CUMULATIVE...
CUMULATIVE...
CUMULATIVE...
ENTRY_TERM
ENTRY_YEAR

# Data Munging/blending/mashup/wrangling

- 3rd party syndicated data
  - One file per year
  - Recast and pivot the data
  - Some questions on format
  - EXTREMELY long titles, more than 128 characters
  - Merged records into single file
- Joins
  - Used left outer joins to maintain completeness
- Bowie state definitions vs syndicated
- Basic star – facts and dimensions

createclearinghouseYR1.txt

```
CREATE TABLE "STUDENTANALYTICS"."CLEARINGHOUSEYR1"
   (    "ENTERING_COHORT_YEAR" NUMBER,
    "REQUESTOR_RETURN_FIELD" VARCHAR2(4000) COLLATE "USING_NLS_COMP",
    "POST_GRAD_YEAR" NUMBER,
    "DATE_RANGE" VARCHAR2(4000) COLLATE "USING_NLS_COMP",
    "School Code/Branch of First Different Two-Year Institution Attended" VARCHAR2(4000) COLLATE "USING_NLS_COMP",
    "Name of First Different Two-Year Institution Attended" VARCHAR2(4000) COLLATE "USING_NLS_COMP",
    "Date of First Enrollment at First Different Two-Year Institution" NUMBER,
    "CIP Code Associated to Enrollment at Different Two-Year Institution" NUMBER
   )  DEFAULT COLLATION "USING_NLS_COMP" SEGMENT CREATION IMMEDIATE
  PCTFREE 10 PCTUSED 40 INITRANS 10 MAXTRANS 255
 COLUMN STORE COMPRESS FOR QUERY HIGH ROW LEVEL LOCKING LOGGING
  STORAGE(INITIAL 65536 NEXT 1048576 MINEXTENTS 1 MAXEXTENTS 2147483645
  PCTINCREASE 0 FREELISTS 1 FREELIST GROUPS 1
  BUFFER_POOL DEFAULT FLASH_CACHE DEFAULT CELL_FLASH_CACHE DEFAULT)
  TABLESPACE "DATA"
as select
"ENTERING_COHORT_YEAR",
    "REQUESTOR_RETURN_FIELD",
FROM STUDENTANALYTICS.CLEARINGHOUSE3
UNPIVOT (
    ("DATE_RANGE",
    "School Code/Branch of First Different Two-Year Institution Attended",
    "Name of First Different Two-Year Institution Attended",
    "Date of First Enrollment at First Different Two-Year Institution",
    "CIP Code Associated to Enrollment at Different Two-Year Institution" )
    FOR "POST_GRAD_YEAR"
    IN (
        ("DATE_RANGE_OF_YEAR_1",
    "Year 1: School Code/Branch of First Different Two-Year Institution Attended" ,
    "Year 1: Name of First Different Two-Year Institution Attended",
    "Year 1: Date of First Enrollment at First Different Two-Year Institution",
    "Year 1: CIP Code Associated to Enrollment at Different Two-Year Institution"
)
 AS 1,
        ("DATE_RANGE_OF_YEAR_2",
    "Year 2: School Code/Branch of First Different Two-Year Institution Attended" ,
    "Year 2: Name of First Different Two-Year Institution Attended",
    "Year 2: Date of First Enrollment at First Different Two-Year Institution",
    "Year 2: CIP Code Associated to Enrollment at Different Two-Year Institution"
)
```

# Visualizing Data Results

- Challenges of large tables
- Topic visualization and layout
- Summary to detail
- Cross-dimensional analysis

- Demo



vlamis™
SOFTWARE SOLUTIONS

# Future Work

- Building a foundation for enrollment analysis
  - ADW Data warehouse project in progress
  - OAC dashboards/canvases design in progress
- Many, many other subject areas
  - Degree analysis
  - Class performance
  - Academic progress
  - Student success

# Q&A



vlamis™
SOFTWARE SOLUTIONS

# Thank You!!


**Tim Vlamis – Vlamis Software Solutions**

tvlamis@vlamis.com



vlamis™
SOFTWARE SOLUTIONS