# CommunityViewer: integrating data for community action



#### Overview

CommunityViewer is an integrated data system created to help San Antonio-area service providers and policymakers understand local needs and respond to those needs in an intelligent and efficient way. The application is intended to allow the integration at the person level of data from multiple collaborating education, housing, justice, case management, and other health and human service organizations and systems in San Antonio, Texas and the surrounding area.

CommunityViewer "stitches together" private agency datasets like those of schools and human service organizations and public datasets like the Census, vital statistics, and hospital discharge data. It geocodes person- and household-level records so that those can be understood in the context of place and neighborhood conditions. It enables trending over time and longitudinal tracking, for, say, the graduation rates of children who attended pre-K versus those who did not. Among other collaborations, CommunityViewer is in use by PaCT (Promise and Choice Together) partner organizations. It currently contains nearly three years of data from SAISD – including about 50 fields and 20 visualized red/green/yellow early warning indicators – as well as demographic, intervention, and attendance data from about 15 out-of-school time (OST) and academic support providers.

CommunityViewer's Public Reports allow anyone with an internet connection and browser to query and download unduplicated aggregate data - never data about individual people - at no cost. The Private Reports, Person View, and Household View are locked down on secure servers and cannot be accessed by the general public. A few specific staff of the partner organizations using CommunityViewer are given the lowest level of access possible for them to perform their job duties. Some can see only grouped data in Private Reports, never information about identifiable people in Person View or Household View.

CommunityViewer is managed by Community Information Now (CI:Now), a local data intermediary serving the San Antonio region. Because agencies do not want to lose control of data they share, the raw data in the system remain the property of the organization that originally collected it and shared it with CommunityViewer. Data-sharing is governed by two kinds of agreements:

- formal written data-sharing agreements between UWSA and the organizations sharing data with CommunityViewer; and
- formal written consent of the person the data is about, his/her parent, guardian, or legal proxy. People control who can see what of their own CommunityViewer data and can change these choices at any time.

#### How CommunityViewer integrates data

CommunityViewer can receive data in one of two ways. The first method, which is lower-risk but higher-cost, is via a physical or virtual "data-to-information" (DTI) server placed behind the data-sharing agency's firewall with a read-only connection to the agency's information system(s). The DTI server converts raw data to information by reading a full copy of the data of interest, converting that raw data into indicators that directly inform a decision or trigger an action, and passing only the indicators and key "person-matching fields" (e.g., name, date of birth, last four digits of social security number) to the central CommunityViewer server. To use school attendance as an example, once a student misses three days in a row, a parent liaison makes a home visit. An OST provider wants to use percent of days missed in the past month as a short-term success measure. In neither case is it important to know which day of the week has been missed, so the DTI reads the daily attendance raw data, but it passes to the central server only the student's person-matching data and the needed indicators: number of school days missed in a row and percent of school days missed in the past month. Because the amount of raw data passed is minimized, the agency is subjected to less risk of a breach of privacy or security.

The second method, which is higher-risk but lower-cost, is direct import of a record-level flat data file (e.g., a csv file in which each row is a unique person) in standardized format into CommunityViewer. This method is the only option in cases where the agency maintains the data of interest only on paper, in a spreadsheet, or in Microsoft Access. It might be the preferred option in cases where the data of interest is needed only once or infrequently, e.g., once per month, and is not needed in or close to real time. The data can still be transformed into information/indicators, but with this method, that transformation happens within the central CommunityViewer database. At present, flat-file data must be imported into CommunityViewer by a database developer with access to the "back end" of the application. If an initiative expects to need a significant number of flat file imports conducted, a dedicated import screen would enable an agency to submit data directly through the front end without the help of a developer.

Regardless of connection method, records with address information are matched with existing records as well as geocoded. New indicators about individual people, families, or households can be created using multiple datasets.

#### How CommunityViewer can be used

#### Planning and measuring progress

Custom aggregate reports run through the **Public Reports** module can provide useful information for identifying and quantifying community issues, planning programs and policy initiatives, and monitoring community change. The two datasets currently available through Public Reports are births and hospital discharges for the entire state of Texas. Both datasets can be queried by year, geography (census tract, zip code, county), sex, age group, race and ethnicity, and key variables in the dataset. Over time these datasets will be updated and others added, including the EDI kinder-readiness index, education indicators, indicators of child well-being, and assisted housing information. Once multiple agencies of the same type are connected to CommunityViewer, important indicators like chronic absenteeism and

STAAR pass rates can be added to Public Reports for query by neighborhood and other ways the data are not typically available.

The private **Custom Reports** module enables custom aggregate queries of the indicators by variables like race/ethnicity, grade level, and LEP (limited English proficiency) status. These reports can be accessed only after secure log-in and with appropriate privileges. The default report presentation (see screenshot following this narrative) is a graphical representation of the data with actual values, and the report can be exported either as a formatted pdf or as an editable spreadsheet. Each report shows all the report parameters so that anyone viewing the report can see at a glance what it includes or excludes.

Available resources haven't allowed this module to be fully developed to include a large number of fields; at present only academic indicators can be queried, and they cannot be cross-tabbed with each other. For example, a user can run a report on percent of high schoolers living in a specific zip code who are chronically absent, and a separate report on percent who received a C or lower in math at the first 9-week report. But one can't look at chronic absenteeism vs. math grade in order to see whether that zip code's students with better attendance are doing better in math compared to those who are chronically absent. But the foundation is there, is user-friendly, and can be built upon as needs and resources present themselves.

More complicated queries like the one decribed above can be done through the secure virtual workstation, which is accessible by only three key staff (one at EPN and two at CI:Now). The virtual workstation is not a good solution for service provider use, but it is appropriate and available for complicated ad-hoc data questions and more exploratory analyses.

#### Early intervention and intervention monitoring

The private Person View, and Household View, available only after secure log-in by designated staff with the user rights to see private identiable information (PII), show the indicators for a specific person, household, or family not living together. In the **Person View**, the current actual value is shown for each indicator and also visualized as red (high risk or poor outcome), yellow (moderate risk or early warning), or green (low risk or good outcome). With one click the user can graph the change in the indicator over time to see the person's trajectory and whether progress is being made after intervention. In addition to looking up an individual person or choosing from a list of people in a selected cohort (e.g., all Bowden Elementary third-graders), users with rights to view PII can also drill down to the individual record from a Custom Report. For example, continuing with the report described above, a service provider might want to know the identities of the students who have good attendance but a poor math grade, as the appropriate intervention for those students might be different from those students with a poor math grade who are also chronically absent.

The main tab of the **Household View** is a dashboard that displays the red-yellow-green status of each household member's relevant indicators, and drill-down functionality enables the user to look at the detail and current value for any given indicator. A separate tab in the Household View captures information about important people who are living outside the household, such as a non-custodial father or a cousin who cares for the children every weekday. Both tabs can be edited to add or remove individuals from the dashboard and to choose which indicator groups are displayed, e.g., secondary education and adult job training but not housing.

For this functionality to work, CommunityViewer must have information about relationships between individual people, e.g., Johnny is the natural son of Josie and the natural brother of Elida. These relationship attributes may not exist in the administrative datasets CommunityViewer connects to and may need to be documented manually by a service provider working directly with that person or household/family. Because families and households change over time, each relationship attribute has a start and end date. If Josie and David divorce in February 2015, that end date is affixed to the husband/wife relationship attribute, but neither the natural mother nor natural father relationship to Johnny and Elida changes.

Likewise, because individual members and whole households move over time, in many cases quite often, address information is never over-written. Each address is tagged with the agency it came from and a start and (if applicable) end date and all current addresses are visible to the user. For query purposes, though, each person must have a single current address, and at present the SAISD address is the default.

#### **Data Control, Privacy, and Security**

The protection of data requires both human and technical/physical safeguards. New users must be approved by a designated authority; because the system is currently in use primarily for EPN and other initiatives administered by United Way, currently only the VP for Grant Implementation at United Way has that authority. That authority will be expanded to additional persons as necessary as the application scales.

The CommunityViewer **User Agreement** attached specifies how the data may and may not be accessed and used. New users must agree to those terms and sign the User Agreement prior to first log-in. The User Agreement also requires that the user receive at least a minimum level of video training in data privacy and FERPA, HIPAA, and Texas Medical Privacy Act (HB300) regulations, and ideally the user receives re-training on key concepts on a recurrent basis. The user also must agree to terms of use at each system log-in. An audit trail logs every user action and is available for review should an issue arise. User rights may be restricted or terminated entirely if these terms are breached.

Three layers of rights and permissions control the use and view of PII in CommunityViewer. First, the person (or parent of the person) whom the data are about must consent to the use of his/her data. Use of the data for aggregate-only reports is opt-out, meaning the data are used unless permission is specifically denied, and data for PII view and any movement of record-level data out of the system are both opt-in, meaning the data are *not* viewed or moved unless permission is specifically granted. The OCR-friendly **Consent Form** is written at the 9<sup>th</sup> grade level in English and in Spanish and is accompanied by an explanation of what CommunityViewer is, how it is being used, and what each consent choice means. (That explanation will likely have to be adapted once some other initiative begins using CommunityViewer.) Consent choices can be changed at any time via any participating provider.

Participants' consent choices can also be communicated verbally in one-on-one settings with service providers. The choices are documented in the CommunityViewer **Consent module** (see screenshot) by the service provider. CommunityViewer then sends a temporary, one-time-use code to the mobile phone the participant specifies for use for that purpose, the participant gives the code to the service provider, and the service provider enters in into CommunityViewer. Consent choices may be entered and saved at any time, but those choices are not "locked" and in effect until the CommunityViewer-generated code is entered. (The code can be re-sent.) This code serves as a proxy for a digital signature and prevents a CommunityViewer user from intentionally or unintentionally making consent choices for a program participant. New functionality being built in spring 2015 is a solution for "mass consent", where a large number of consent forms are gathered in the course of school or program enrollment and the provider and participant are not able to sit together one-on-one with the time for a conversation about data use and consent and the exchange of the lock code.

Second, the agencies wishing to share data with each other must negotiate and execute some kind of formal Data-Sharing Agreement outlining the terms of that sharing. A very over-simplified metaphor is that CommunityViewer provides the phone line, but the agencies must still call each other in order to talk. Only the agency that owns the data can grant permission for that data to be shared with another agency via CommunityViewer. Depending upon agency needs and preferences, the agreement may take the form of brief language incorporated into a basic service contract between the two agencies, a general data-sharing MOU, a detailed technical agreement, or an extensive formal document vetted by attorneys on both sides. The agreement may be a hybrid, e.g., a high-level formal MOU approved at the board level that codifies the intent to share data, and a more detailed and technical agreement executed at the staff level that specifies what and how data will be shared. One example of boilerplate for that technical agreement is included here as an appendix. (If private health information is being shared, a separate HIPAA Business Associate Agreement (BAA) must be executed in order to comply with the Texas Medical Privacy Act, even if neither agency is a covered entity under the federal HIPAA rule.)

Third, the **User Rights** module controls which datasets a user has access to. A user is by definition allowed to see any data that comes from his or her own employer, although s/he might not have rights to view PII. But because a user should see only the data necessary to perform his/her job and multiple initiatives are and will be sharing data through CommunityViewer, most users will have access only to the data from agencies it is working with, and should never be able to see PII for people they are not working directly with. Further, a user from an OST agency or school district might be able to see Generic Nonprofit's education information for participants they both work with, but not any housing or health information.

#### **Technical and physical safeguards**

Regulatory Compliance. Privacy of all person-level records is maintained by utilizing two-factor authentication and the encryption (Advanced Encryption Standard [AES] 256) of all personally identifiable information (PII) and electronic personal health information (ePHI) while traveling in the public network and at rest within the Oracle database. PII and ePHI are not synonymous terms, but the risks and safeguards are similar and they will be treated together here except for provisions specific to HIPAA regulations.

Automation tools are used to identify assets in the Virtual Data Center (VDC) and find any vulnerabilities, misconfigurations and malware exposure that may put ePHI at risk. Contextual information into the level of risk posed by each vulnerability will help prioritize remediation and mitigation to ensure PII/ePHI is truly secure. HIPAA/HITECH compliance is maintained utilizing these "always on" tools. Internal penetration testing is also conducted to verify all configurations in the VDC.

#### The managed hosting provider can:

- Detect PII/ePHI data in environment.
- Get top-down visibility of risk to assets and business operations, enabling us to organize and prioritize assets and quickly focus on the items that pose the greatest risk.
- Get a clear map of the real risk posed to ePHI by the identified vulnerabilities across the VDC landscape.
- Combined with Common Vulnerability Scoring System (CVSS) base scores, temporal scoring, environmental considerations (e.g., any mitigating controls in place), and asset criticality for risk classification.
- Take inventory of systems, services, and installed applications using the latest fingerprinting technologies.
- Detect the presence of unauthorized software on information systems and notify designated officials through alerts on an automated mechanism.

- Perform comprehensive unified vulnerability scanning of all vital systems including networks, operating systems, web applications, databases, enterprise applications, and custom applications.
- Generate detailed reports combined with role-based access controls
- Compare the results of vulnerability scans over time to determine trends in information system vulnerabilities through an automated mechanism.
- Audit users and groups on systems.
- Set-up automated monitoring access controls (including adherence to policies for role-based access) to validate enforcement of access restrictions.
- Test the efficiency of access control systems and policies

The host's data center is ruled under the following FISMA policies:

- Access Control
- Awareness and Training
- Audit Accountability
- Contingency Plans
- Incident Response Policy
- Risk Assessment Policy
- Physical and Environmental Policy
- Planning Policy and Procedures
- Configuration Management Plan
- Information Security Policy
- Identification and Authentication Policy
- Systems Communication Policy and Procedures

These policies along with a system security plan are available upon request.

Administrative Safeguards. The development and hosting provider requires that all personnel complete all training requirements outlined by the training manager of the contract, consistent with the methodology of National Institute of Standards and Technology (NIST) Special Publication 800-50, Building an Information Technology Security Awareness and Training Program, which applies to all employees as well as remote researchers and collaborators working on its projects. The component must include awareness activities, basics and literacy training activities for general users, and role-based training for specialized users. In accordance with the Code of Federal Regulations, Title 5, Part 930, Subpart C, Section 930.301 (a)(1), each user is required to engage in annual basics and refresher training to sustain such access. In addition, training is conducted on HIPAA Awareness, HIPAA Privacy Rule, which applies to safeguarding Protected Health Information (PHI) in oral or written form and HIPAA Security Rule, which apply to electronically stored or transmitted PHI and technical protections for

electronic Protected Health Information (ePHI).

*Physical Safeguards.* Public access to the data center is forbidden. Only personnel with the appropriate clearance are allowed in the server areas of data center, and the two-person rule always applies. Access to the data center is monitored by a security system and security cameras and controlled by two independent security doors with both cypher and magnetic card locks. Two independent monitoring companies (ADT and Protection One) actively monitor the data center at all times.

Technical Safeguards. CommunityViewer has been designed as a redundant system with 24/7 availability with 99.9% up time. The system contains two CISCO ASA firewall systems, one Proxy Firewall, redundant application servers on a de-militarized zone (DMZ), and a redundant Oracle database. Compliance patching has been done utilizing IBM software BigFix that constantly monitors for FISMA compliance. Following a recent purchase, we are in the process of installing and configuring IBM VMware vCloud Suite Enterprise with Health Insurance Portability & Accountability Act (HIPAA) and Health Information Technology for Economic & Clinical Health (HITECH).

Compliance and Configuration Management. The Public Reports interface and aggregate data queried through it are freely accessible by anyone with an internet connection and web browser. (Data is suppressed where cell size is lower than 30.) Login to the secure area (Private Reports and the Person Viewer person- and family-level dashboard) person-level record requires two-part authentication and user acceptance of terms and penalties for violating those terms. (Two-part authentication requires knowing the username and password and having a one-time use code or "token" generated via an app called WiKID.) All system use and edits are logged.

The development and hosting vendor utilizes programming controls that obscure the feedback of authentication information. All information traveling via the public network is encrypted using Federal Information Processing Standard (FIPS) compliant encryption algorithms.

Breach Notification. In general, HIPAA sets the highest legislated privacy and security standards and those regulations are much farther-reaching than FERPA. But criteria and procedures similar to those below are used to define and handle an "incident" involving data contributed by agencies not considered to be a Covered Entity under HIPAA regulations.

An incident will be considered any physical, technical or personal activity or event that increases the (HIPAA) Covered Entity's risk to inappropriate or unauthorized use or disclosure of PHI or causes the Covered Entity to be considered non-compliant with the Administrative Simplification provisions of HIPAA/HITECH as determined by the Department of Health and Human Services. Notification will be made by Business Associate to Covered Entity by telephone or secure fax of any HIPAA/HITECH

Electronic Transactions and Code Sets, Privacy, Security or Standard Identifier Incident, or Use or Disclosure of PHI not provided for by this BAA.

A written report of the incident, submitted to Covered Entity within ten (10) business days after initial notification, will document specifics surrounding the incident, what mitigation procedures were implemented to lessen the impact of the incident and what processes have been established to prevent the incident from occurring in the future (reasonable and appropriate safeguards). This report should be documented as a letter.

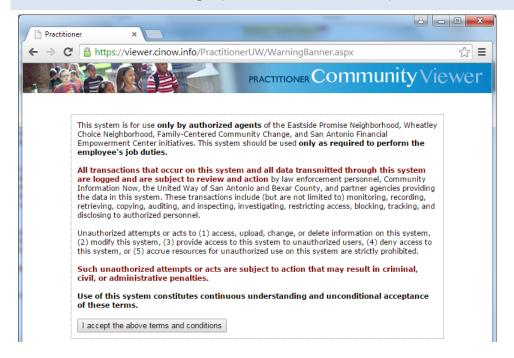
Secure Managed Hosting. CommunityViewer could be hosted anywhere with acceptable secure managed hosting capacity. The current secure managed hosting provider's data center is specifically designed to receive and process, display, store, and transmit data in an ultra-secure environment. To be able to host data and applications for federal customers, the vendor has met all Federal Information Security Management Act (FISMA) requirements and has earned U.S. Department of Defense Information Assurance Certification and Accreditation Process (DIACAP) certifications. Its systems are Health Insurance Portability and Accountability Act (HIPAA) compliant, Health Information Technology for Economic and Clinical Health (HITECH) compliant, and Family Educational Rights and Privacy Act (FERPA) compliant.

The secure managed hosting vendor has the processes in place to handle large databases up to 42 terabytes. The web server utilizes an X.509 PKI Server Certificate and has 128-bit Transport Layer Security (TLS) version 1.1 (or SSL version 3.0) enabled at all times. All connections are via https (128-bit encrypted SSL) over port 443. All encryption algorithms are National Institute of Standards and Technology (NIST) / Federal Information Processing Standards (FIPS) certified.

The secure managed hosting vendor assures 99.99% uptime, offers High Availability and Fault Tolerance, and the system is monitored 24/7. If a total utility power outage occurs, all data centers' power systems are designed to run uninterrupted, with every server receiving conditioned Uninterruptible Power Supply (UPS). If an extended utility power outage occurs, the vendor's routinely tested, on-site generator can run indefinitely.

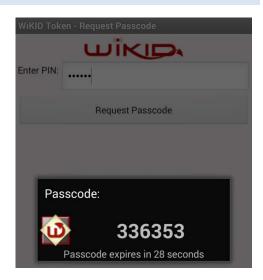
#### **Selected screenshots**

### Terms and Conditions at login (can be revised as needed)



### Two-part authentication required at login



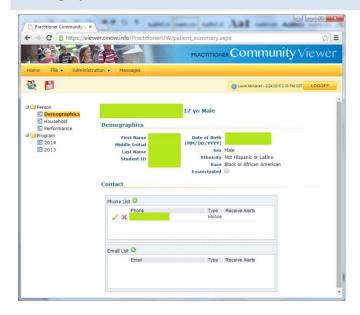


# **User Management module**

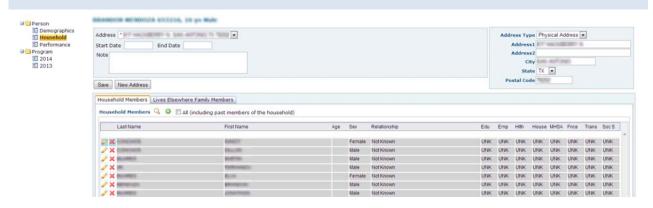




### **Demographics View**

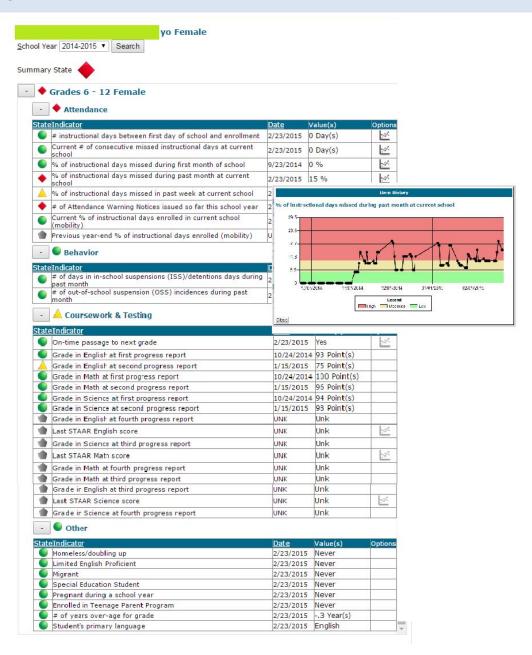


#### **Household View**



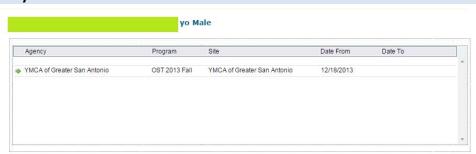
#### **Performance View**





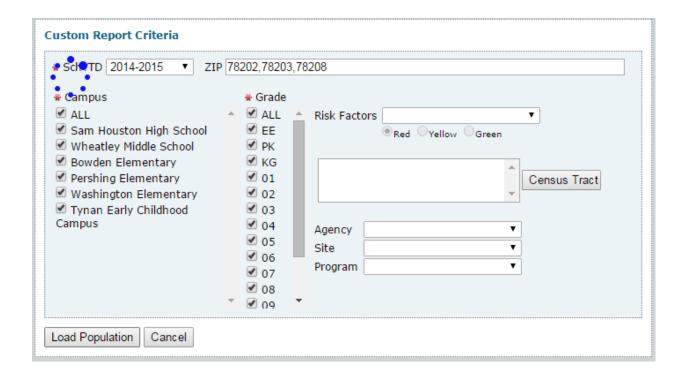
### **Program Summary View**







### **Custom Reports: Select criteria**

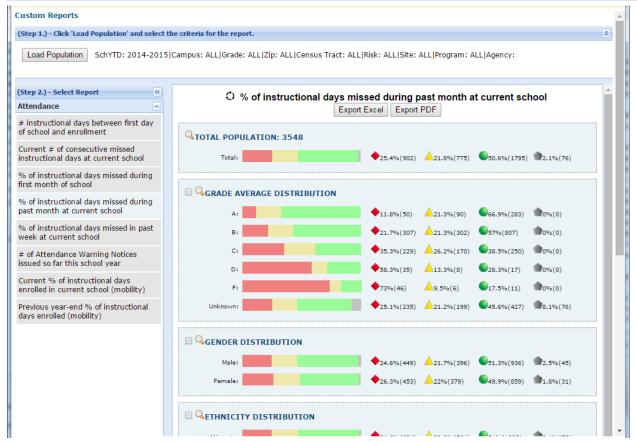


# Custom Reports: Sample output: students from three selected zip codes only

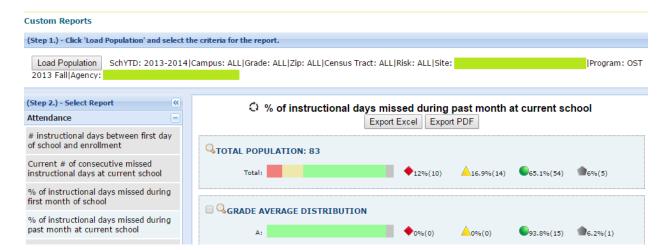


# Custom Reports: Sample output: students from all zip codes

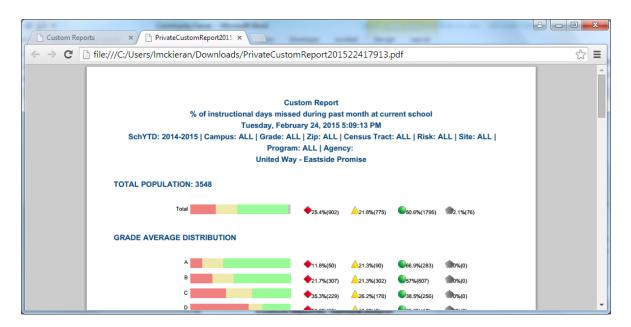
(comparison shows students from the three-zip code area are less likely than the student population as a whole to have been chronically absent in the past month)



#### Custom Reports: Sample output: students served by OST provider (redacted), fall 2013



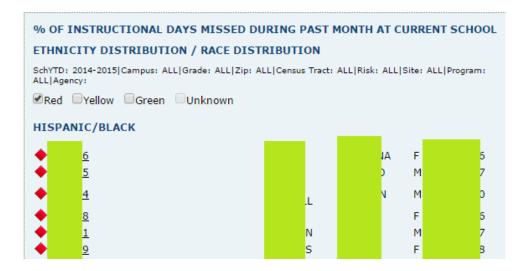
# **Custom Reports: Sample output exported to pdf**



#### Custom Reports: Sample output crosstabbed by ethnicity and race



# Custom Reports: Drill-down (with PII rights) from aggregate report to list of individual Hispanic/Black students who are chronically absent



# Custom Reports: Drill-down (with PII rights) from above list to the first student listed

