UCAS Database of Schools and Colleges

DG/14/006

This paper was produced by Alison Charles from The University of Manchester. Wendy Webster, University of Dundee will talk through it at the Data Group meeting.

Circulation   UCAS Undergraduate Advisory Group

From   Alison Charles, Student Admissions Manager, The University of Manchester

Date/Version   February 2014

The University of Manchester propose that:

a. Improvements are made to the UCAS database of schools and colleges (hereby referred to as the UCAS schools database);
b. As part of these improvements, UCAS match the school codes to the school and college codes assigned by each of the four UK administrations (hereby referred to as Education codes);
c. As part of these improvements, UCAS consider whether incorporating the UKPRN (UK Provider Reference Number)to the UCAS schools database and the school performance data for each country would resolve the issues outlined in this paper. UCAS to carry this forward with each of the four UK administrations if deemed appropriate.

1. **Contextual Data**

The University of Manchester Student Admissions Office collates contextual information for our admissions staff in academic Schools for a full and rounded view of applicant achievement and potential. A candidate is flagged if they meet a specific combination of geo-demographic, educational and local authority care indicators.

We collect publicly available datasets, including school and college performance data at Level 2 and Level 3 (or country equivalent) from The Department for Education (DfE), The Welsh Assembly Government, The Scottish Government, and The Department of Education Northern Ireland (DENI).A candidate meets our educational indicator when their educational establishment performs below the national average across three years (where data is available).

For more details please see www.manchester.ac.uk/contextualdata.

The University of Manchester annually reviews the basket of contextual data provided by UCAS. We do not currently use this data for the following reasons:

- Whilst the contextual data provided by UCAS is matched to individual applicants, it is not put into context; for example national averages are not provided for a comparison to be made.

- By contrast the school performance data we collect is compared to its corresponding national average using z-score calculations and the resultant calculation is displayed in our student system with the appropriate flag assigned. The flag provides a clear indication to the admissions officer of the need for additional consideration.

- As well as providing school performance data matched to individual applicants, UCAS also supply raw data in a dataset. However this is not provided by country and therefore national averages cannot be applied. This is particularly important given the differences in education systems at level 2 and level 3 particularly in Scotland.

Therefore, in order to ascertain the performance of the school or college an applicant has attended, we map three years of performance data for individual schools and colleges (as provided by each UK administration) against the UCAS schools database. When we find establishments that have performance data but are missing from the UCAS schools database, we request new UCAS school codes from UCAS. There are fundamental issues with the quality of the UCAS schools database and also the DfE school performance data (for instance DfE codes are recycled when schools close). These are outlined in the appendix. The combination of these issues means that mapping schools and colleges across three years of performance data is not a reliable process.

We have committed a significant amount of staff time to match the data and then to manually check it. Our Systems Development Officer has committed several months to developing and annually improving an SQL database to match the data. In addition to this, he spends approximately three weeks each year matching the revised data for each country and our Student Admissions Officer spends approximately three weeks each year checking the matched data before and after it has been uploaded to our student system.

2. **Planning Support**

The University of Manchester Planning Support Office (PSO) are required to submit information about the previous education of our students via the Previous Institution field in the annual HESA return. This field is currently populated with either UCAS or DfE codes, and so incurs an identical mapping problem when the PSO receive freedom of information requests or internal enquiries regarding individual students. (For example, a student's school codes can map to multiple rows of the UCAS schools database, and these different rows can have different values for school type.)

3. **HESA**

   The Higher Education Statistics Agency have requested a non-PDF version of our published Educational Indicators document (which lists UK schools and colleges, including both UCAS school code and Education code), and/or information on how we have mapped these. This is to help them with data requests regarding the Previous Institution field in the HESA returns. This further strengthens our case for the quality of this data to be improved and for it to be matched nationally.

   For students commencing higher education in 2014, HESA will use the UKPRN(UK Provider Reference Number) for the Previous Institution field.

4. **SPA**

   The University of Manchester has contacted Janet Graham, Director of SPA (Supporting Professionalism in Admissions) as contextual data and related good practice are a major objective for SPA. SPA is keen to understand the issues that institutions are having in matching data in order to support this request, and has offered if appropriate to try and broker a meeting between UCAS, national departments for education, HESA and any other relevant stake holders to discuss how best resolve the matching and related issues. SPA supports our decision to raise the request at the UCAS UG Advisory Group as to resolve this through UCAS as a national shared service would be the best way forward. Resolution at a national level would be better than one institution providing its matched data and methods to HESA.

5. **Summary**

   The University of Manchester propose that:

   a. Improvements are made to the UCAS database of schools and colleges;
   b. As part of these improvements, UCAS match the UCAS school codes to the school and college codes assigned by each of the four UK administrations.
   c. As part of these improvements, UCAS consider whether incorporating the UKPRN to the UCAS schools database and the school performance data for each country would resolve the issues outlined in this paper. UCAS to carry this forward with each of the four UK administrations if deemed appropriate.

   This will eliminate the duplication of work that takes place in Admission and Planning Support teams in HEIs across the country. This proposal is timely given the issues HESA are experiencing with the Previous Institution field and the changes to the DfE performance tables.

Appendix

Specific issues with the UCAS schools database and DfE school performance data.

At The University of Manchester, we have significantly changed the process of collecting the data from UCAS over the last 3 years. Originally the data was collected from the *cv school table* at UCAS, which was then matched using an Access database to any known relationships with the DfE codes. There was little confidence in the up-to-date nature of the UCAS data in this process (circa 2010) as the DfE updated their codes on an annual basis and as there was little to no communication with UCAS about these changes, a large proportion of the matched data was either incorrect or out of date within a year – primarily with English institutions.

Any changes to institutions, either by merging or reuse of existing DfE codes, meant that the dataset needed to be verified manually before confirming the relationship between the two codes – this took a long time in the case of England, as there are over 6000 institutions in the published data each year. In addition to this, as new UCAS codes can be created for institutions that may not have changed their DfE code, you have a one to many relationship that needs to be checked and verified each year if you wish to show an institution's historical data (i.e. 6 known UCAS codes can relate to the 1 DfE code). As the current MERGEDWITH column at UCAS only provides the one last linked relationship, this can prove challenging for a database to automatically match results in this way.

Additionally, as the data held in the Previous Institution Name column at UCAS may differ slightly from the naming convention listed at the DfE, the only way to check that schools were linked correctly was to go through a large dataset and manually verify the relationship.

Since 2011 we have tried to communicate any new or unmatched DfE codes to UCAS so that they can update their database and where necessary provide us with new UCAS codes. We then manage this list from a central database containing the linked relationship between these many different codes – supplementing this list with any new codes/relationships each year as new data is produced.