

Introduction

This Annex lists the data-fields available from the Airwave Study Tissue Bank employments export. The reader should consult the main data-dictionary for a description of the screening protocol itself, explanations of formatting conventions, missing values, and other technical information.

Document Configuration

Title	Data Dictionary, Annex B
Subject	Employments Export
Version	2.1
Author	Heard, Andy H, Database Manager at Imperial College London.
Published	14/08/2026
Filename	annex-b.docx

Changes since Previous Version

This is a new export.

Employments

Each record is a single observation of basic details about a participant's employment circumstances. Many participants will have several records over time.

We have taken care to merge descriptions of job-rank and job-role where this information might, when taken together with information in other datasets, risk individuals being made identifiable. Researchers are nevertheless reminded of their obligation not to attempt to carry out identification of individuals from data provided to them.

These data have been derived from questionnaires, information gathered at the health screen, and datasets provided by Police services' Human Resources departments.

Methods of Clustering

The four cluster variables and this section of documentation were derived from material supplied in August 2022 by Dr Ioannis Basinas, a researcher from the Centre for Occupational and Environmental Health at Manchester University.

Clustering of police forces is based on analysis of official data as published by the ONS and police forces. Specific parameters considered include a) the level of crime in the police force area by specific type, b) the density of the population in the police force area and c) the index of multiple deprivation. Data were extracted by the ONS database - i.e. crime in England and Wales statistics, English and Wales indices of deprivation and census 2011 for density.

Once pooled in a common database elaborated in MS Access the above data were analysed to establish the related clusters. Cluster analysis on the basis of K-means (flat clustering) is implemented. Analysis has been performed in SAS v9.4. Prior to clustering correlations between variables were examined and all variables were standardised to a mean of zero and standard deviation of one (converted to z-scores). As strong correlations between the IMD scores and the crime rates measures were present the IMD scores were excluded from further analysis. K-means Clustering was performed using the PROC FASTCLUS procedure. Number of clusters tested in the

process varied between 3 and 10 with performance of the clustering process was evaluated by standard approaches including the Pseudo F and Cubic Clustering Criterion.

Label	Data Type	Description
part_id	NUMBER (7)	Pseudonymous identifier that is unique for each participant. Because the screening extract includes follow-up visits, some part_id values are associated with two or more barcodes.
region	STRING	Geographical region in which their force is located at the time they provided the information.
observed	DATE	When the information was collected.
job_rank	STRING	A merged description of the participant's rank within their police force.
job_role	STRING	Description of the job-role undertaken.
cluster_victim	NUMBER(1)	Cluster of police forces based on population density and rates of victim based crime in the police force area. Rates of crime per 1000 population used. Victim based crimes include violent crimes against the person (i.e. homicide, violence without injury, stalking and harassment, death or serious injury) sexual offences, Robbery, and Theft offences.
cluster_violence	NUMBER(1)	Cluster of police forces based on population density and rates of violent crime against the person in the police force area. Rates of crime per 1000 population used. Violent crime against the person includes homicide, violence without injury, stalking and harassment, death or serious injury, sexual offences and Robbery
cluster_Other	NUMBER(1)	Cluster of police forces based on population density and rates of other types of crime against society in the police force area. Rates of crime per 1000 population used. Other crimes against society include drug offences, possession of weapons offences, public order offences and miscellaneous crimes offences.
cluster_total	NUMBER(1)	Cluster of police forces based on population density and rates of total crime in the police force area. Rates of crime per 1000 population used. Total crime includes all types of crime including victim based and society based.