

ICPSR 37698

**Dunham's Data: Katherine Dunham
and Digital Methods for Dance
Historical Inquiry, Everyday
Itinerary, 1947-1960**

Harmony Bench
Ohio State University

Kate Elswit
*Royal Central School of Speech and Drama
(Great Britain)*

P.I. User Guide

Inter-university Consortium for
Political and Social Research
P.O. Box 1248
Ann Arbor, Michigan 48106
www.icpsr.umich.edu

Terms of Use

The terms of use for this study can be found at:
<http://www.icpsr.umich.edu/web/ICPSR/studies/37698/terms>

Information about Copyrighted Content

Some instruments administered for studies archived with ICPSR may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

NOTICE

WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

This User Guide is intended to accompany the Everyday Itinerary Dataset, 1947-60 from *Dunham's Data: Katherine Dunham and Digital Methods for Dance Historical Inquiry*. The first version, encompassing 1950-53, was released August 2020 as the first dataset in our series. The second version, 1947-1960, will be released in September 2022.

The Dunham's Data project is funded by the UK Arts and Humanities Research Council (AHRC AH/R012989/1, 2018-2022), under the direction of Kate Elswit (PI, University of London, Royal Central School of Speech and Drama) and Harmony Bench (CI, The Ohio State University). Through this project, we explore the kinds of questions and problems that make the analysis and visualization of data meaningful for dance history. We do so through the case study of choreographer Katherine Dunham, manually cataloging a daily itinerary of Dunham's touring and travel from the 1930s-60s, the dancers, drummers, and singers in her employ during that time, and the repertory they performed. These curated datasets provide new means to understand the relationships between thousands of locations, and hundreds of performers and pieces across the decades of Dunham's career, and ultimately elaborate how movement moves.

The Everyday Itinerary Dataset, 1947-1960, encompasses Katherine Dunham's daily locations, travel, and performances every day over fourteen years of her most substantial period of consistent international touring. Over this time, Dunham's personal and professional travels took her to 194 unique cities over 429 trips, on every continent but Antarctica. This dataset tracks geographic location (98% of 5110 days); and, less comprehensively, the accommodation in which Dunham stayed each night; the theatres, nightclubs, television studios, and other places she and the company performed; the modes of transportation used when travel occurred; additional transit cities through which she passed; and whether or not Dunham was likely to be in rehearsals or giving public performances.

We have chosen this user guide model over a more conventional codebook because the data is relational; for example, venues in different cities may share the same name, and when these are dissociated from city names, any resulting data analysis is misrepresentative. A code book therefore cannot accurately capture the crucial inter-relationships among the elements our datasets document. In the primary dataset, we maintain Dunham's level of descriptive specificity (eg: Harlem is different than New York City), but the supplemental dataset "Everyday Itinerary 1947-60 Grouped Cities Data" groups locations by proximity. The primary dataset is curated by Harmony Bench and Kate Elswit, with supplementary datasets created by Antonio Jiménez-Mavillard on the basis of that data (see below).

The Everyday Itinerary was created, audited, and released in multiple phases. Data for all versions were manually curated from undigitized archival sources including travel documents, personal and professional correspondence, receipts, desk calendars and log books, payroll and other company documents, newspaper clippings, and so on. The Everyday Itinerary Dataset, 1950-53 was created first as a test-case. After the dataset's initial curation, the entire dataset was audited by thoroughly reviewing the same archival documents in reverse chronological order to correct for plans that did not materialize. As a touring artist, Dunham routinely kept

multiple possibilities for engagements open, and even seemingly solid plans were sometimes changed at the last minute. Additional evidence was sought from dance criticism and event notices available through newspaper databases. Both source materials and decision-making processes have been documented to facilitate our own data curation and auditing process, as well as to aid future scholars in navigating these materials. This same process was then brought to the 1947-60 expansion, and to the future 1937-62 version.

For more on this dataset and its analysis, see [“Katherine Dunham’s Global Method and the Embodied Politics of Dance’s Everyday”](#) in Theatre Survey. The data can be interactively explored via an 3D globe ([Katherine Dunham’s Global Travel, 1947-60 \(Interactive Space-Time Mapping\)](#)) or on a timeline ([Interactive Timeline of Katherine Dunham’s Travel, 1947-1960](#)).

For more on the full series of datasets, see [“Dance History and Digital Humanities Meet at the Archives: An Interim Project Report on Dunham’s Data” \(2020\)](#), which is part of the deposit under “DS0 Study-Level Files.”

DATE

The sheet is sorted by date, day by day from January 1st 1947 to December 31st 60, sorted chronologically. The itinerary represents the whereabouts of Katherine Dunham on both personal and professional travels.

CITY 1 & COUNTRY 1

These two columns identify the first location we have in a given day for Katherine Dunham’s itinerary. Most often she spent the previous night in the same location, but this is only confirmed if that location appears in either CITY/COUNTRY 1 or CITY/COUNTRY 2 of the previous day. These two columns should be treated together in order to disambiguate city names that appear in more than one country. CITY and COUNTRY are designated by the names and boundaries used at the time of travel (for example, West Germany and Yugoslavia). To aid in consistent data linking, city names have been Anglicized (for example, Florence and Turin rather than Firenze and Torino). Researchers will note that there is not a separate column for state/province/regional identifier. However, we have retained Puerto Rico as an independent entity in the data rather than subsume it into the United States, of which it is a territory.

CITY 2 & COUNTRY 2

These two columns are only used where we have evidence that Dunham ended her day in a different place than she began it. Any travel after midnight is associated with CITY/COUNTRY 1 the following day. In general, the last city location of a given day is where any associated PERFORMANCE VENUE and HOTEL/ADDRESS is located, although there are exceptions, particularly when a PERFORMANCE VENUE 2 is noted. As with CITY/COUNTRY 1, the CITY/COUNTRY 2 columns should be treated together in order to disambiguate cities with the same name in different countries, and city and country names and political boundaries reflect those used at the time of travel.

HOTEL/ADDRESS

This column contains any information that we have regarding where Dunham stayed overnight on the evening of that day. In general, this refers to hotels or to the addresses of private apartments or houses. Where we have used more descriptive information from the archives, that is noted with quotations. As with PERFORMANCE VENUE, this column should only be treated in conjunction with CITY.

NOTE: there is less archival evidence available for this column, and Dunham often moved accommodations within a city, so dates and duration should be treated with caution.

PERFORMANCE

This column designates whether there is evidence that one or more performances occurred on a given date: (y)es, (n)o, or a blank that means we do not have sufficient evidence to make an assessment. These are often filled down during longer runs, although where we have evidence that a city prohibited Sunday performances, those are marked as no. Only performances are marked yes; non-performance time spent in theatres, including load in/out or dress rehearsals are designated no.

PERFORMANCE VENUE

This column names the primary venue in which Dunham was working on a given day. This is filled out for both performance and non-performance (e.g. rehearsal) time. To ensure accuracy, PERFORMANCE VENUES need to be associated with the row's CITY/COUNTRY, since many performance venues around the world share the same name, and they should not be conflated. Researchers are advised that this column may at times represent colloquial rather than formal references to venue names.

VENUE TYPE

This column captures the variety of spaces in which Dunham performed, from high art venues to cabarets. We have captured these in the column venue type, which designates the kind of performance venue in which Dunham appeared. The possibilities are concert hall (c); screen (including television and film shoots) (s); nightclub, casino, hotel (n); and alternative venues (including high schools and army bases) (a) . Where these categories are fuzzy, we have used the performance programs to guide us — for example, many venues at the time were current or former cinemas, but if Dunham's performances there followed the distinct format of a concert hall performance, that is how we categorize the venue. Due to the local specificity of performance venues, perceptions of their relative prestige, regular programming and clientele, and how these changed over time, VENUE TYPE should be taken as indicative rather than authoritative. The column VENUE TYPE is also used in the Repertory Dataset to represent the particular types of venues in which individual works were performed.

PERFORMANCE VENUE 2

Dunham often performed multiple times at the same venue in one night, however, this column captures only the names of additional venues in which Dunham appeared, when she gave performances at more than one venue in a single day. Like PERFORMANCE VENUE,

PERFORMANCE VENUE 2 needs to be associated with the row's relevant one or more entries for CITY/COUNTRY, since many performance venues around the world share the same name, and they should not be conflated.

VENUE TYPE 2

Like the column VENUE TYPE, VENUE TYPE 2 designates the type of PERFORMANCE VENUE 2 in which Dunham appeared. The possibilities are concert hall (c); screen, including television and film shoots (s); nightclub, casino, hotel (n); and alternative venues, including high schools and army bases (a) .

NOTES

In the notes column, we have provided explanatory notes to assist researchers as they go through the Everyday Itinerary dataset. Such notes clarify the types and dates of documents we have consulted, point out conflicts in the archives, and generally offer a synopsis of decision-making underlying each datapoint. Where notes would otherwise be repeated across multiple consecutive days, notes appear on the first and last relevant days, and intermittently as appropriate.

SOURCE

This column indicates the sources for materials that provide evidence for information contained in a given row. Most are from the special collections at Southern Illinois University at Carbondale (SIU), Missouri Historical Society (MHS), and Library of Congress (LOC). We have indicated both boxes and folders where we have found supporting evidence, both to document our own process and to aid future researchers. We also note newspaper articles, official government documents, transit timetables, and other relevant materials.

SUPPLEMENTARY DATASETS

Everyday Itinerary 1947-60 Grouped Cities Data: This is an adaptation of the main dataset "Everyday Itinerary Dataset, 1947-60." The only difference with respect to the original version is that locations are grouped by proximity. If the geodesic distance between two or more locations is less than a specific threshold, we grouped them in one single location with a common name. This threshold was calculated experimentally for this particular dataset and set to 15 km / 9.32 miles, in order to account for distance, terrain, and geopolitical considerations. The resulting grouped areas are: Port-au-Prince and Petion-Ville under the common denomination "Port-au-Prince." Los Angeles, Hollywood, and Beverly Hills under the common denomination "Los Angeles area."

Stay-Lengths 1947-60 Data: This dataset is based on Everyday Itinerary Dataset, 1947-60. It groups every single day of Dunham's itinerary into stays. A stay is a visit to a place and its length is calculated as the number of consecutive nights spent in that place. Basic columns: CITY & COUNTRY: City visited and its country. START_DATE: Arrival and departures dates for this stay in the city. N_ROWS: Number of consecutive rows (each row represents a day) in

which this city is the value for the column CITY1 on the previous Everyday Itinerary 1947-60 Grouped Cities Data. LAST_MOMENT: Last likely moment spent in the city with respect to the last night spent. The two possible values are: "probably that night" if CITY2 is blank on the last row. We assume they stay overnight if we have no certainty that they travelled. In this case, the last row corresponds to the last night spent in the city. "The next morning" if the value for CITY2 is different from CITY1. The second-to-last row corresponds to the last night spent in the city, while the last row indicates they travelled the following day in the morning. Due to the uncertainty created by the information (or lack of it) contained on CITY2, the stay length ranges between two values: MIN_NIGHTS and MAX_NIGHTS. MIN_NIGHTS is the certain number of nights spent in the city, regardless of the value for CITY2, which is always equal to either N_ROWS (when CITY2 is not blank on the last row and the LAST_MOMENT was "the next morning") or N_ROWS minus 1 (when CITY2 is blank on the last row and the LAST_MOMENT was "probably that night," meaning they travelled in the morning and did not spend the night on their last day). MAX_NIGHTS is the likely number of nights spent in the city, under the assumption that a blank value for CITY2 on the last row means they probably stayed that night. Other columns: LATITUDE & LONGITUDE: Geographic coordinates of the city.